

Lecture Notes in Statistical Mechanics

Igor Vilfan

The Abdus Salam ICTP, Trieste, Italy
and
The J. Stefan Institute, Ljubljana, Slovenia

Trieste, Autumn 2002

Preface

The present Lecture Notes in Statistical Mechanics were written for the students of the ICTP Diploma Course in Condensed Matter Physics at the Abdus Salam ICTP in Trieste, Italy.

The lectures cover classical and quantum statistical mechanics with some emphasis on classical spin systems. I give also an introduction to Bose condensation and superfluidity but I do not discuss phenomena specific to Fermi particles, being covered by other lecturers.

I.V.

Contents

Preface	iii
Tables	ix
1 Foundations	1
1.1 Introduction	1
1.2 Ensembles in Statistical Mechanics	4
1.2.1 Microcanonical Ensemble and the Entropy (Isolated Systems)	4
1.2.2 Canonical Ensemble and the Free Energy (Systems at Fixed Temperature)	6
1.2.3 Grand Canonical Ensemble (Open systems)	10
1.3 Analogy Between Fluids and Magnetic Systems	12
1.4 Fluctuations	13
1.5 Phase Transitions	17
1.5.1 Order Parameter	18
1.5.2 Critical Point	19
1.5.3 Critical Exponents	21
2 Classical Models	25
2.1 Real Gases	26
2.2 Ising Model	30
2.2.1 Lattice Gas	31
2.2.2 Binary Alloys	32
2.2.3 Ising model in $d = 1$	33
2.2.4 Ising Model in $d = 2$	38
2.3 Potts and Related Models	40

I. Vilfan	Statistical Mechanics	Contents
2.4	Two-dimensional xy Model	42
2.5	Spin Glasses	46
2.5.1	Some Physical Properties of Spin Glasses	49
3	Quantum Models	53
3.1	Quantum Statistics	53
3.1.1	The Density Matrix	54
3.1.2	Ensembles in Quantum Statistical Mechanics	55
3.2	Bose Systems	58
3.2.1	Ideal Bose Gas	58
3.2.2	Bose Einstein Condensation	61
3.3	Liquid ^4He	67
3.3.1	Superfluidity	68
3.3.2	Vortex Excitations	69
3.3.3	Order Parameter	70
4	Methods of Statistical Mechanics	73
4.1	Mean-Field Theories	73
4.1.1	Weiss Molecular Field of an Ising System	74
4.1.2	Bragg-Williams Approximation	75
4.1.3	Landau Theory of Continuous Phase Transitions	79
4.1.4	Critical Exponents in the Landau Theory	85
4.1.5	Short-Range Order and Fluctuations	86
4.1.6	Validity of The Mean-field Theory	88
4.2	Monte Carlo Simulations	90
4.2.1	Importance Sampling	90
4.2.2	Metropolis Algorithm	91
4.2.3	Practical Details and Data Analysis	91
4.3	The Renormalization Group Methods	96
4.3.1	A Trivial Example: The $d = 1$ Ising Model	101
4.3.2	RG Transformations on the $d = 2$ Ising Model	104
4.3.3	General Properties of the RG Transformations	107
4.4	Momentum-Space Renormalization Group	117
4.4.1	The Gaussian Model	118
4.4.2	The Landau-Wilson Model	121
4.4.3	Critical Exponents	123

I. Vilfan	Statistical Mechanics	Contents
		<hr/>
	Further Reading	129
	Appendices	131
A	Distribution Functions	131
B	Maxwell Relations	133
C	Basic Thermodynamic Relations of Magnetic Systems Revisited	135

Tables

Table 1: Ensembles in statistical mechanics

Ensemble	Partition function	Thermodynamic potential
Microcanonical	$\Delta\Omega(E, V, N)$	$S(E, V, N) = k_B \ln \Delta\Omega(E, V, N)$
Canonical	$Z(T, V, N)$ $= \sum_i \mathbf{e}^{-E_i(V, N)/k_B T}$ $= \int \Delta\Omega(E, V, N) \mathbf{e}^{-E(V, N)/k_B T} dE$ $Y(T, p, N)$	$F(T, V, N) = -k_B T \ln Z(T, V, N)$ $G(T, p, N) = -k_B T \ln Y(T, p, \mu)$
Grand canonical	$\Xi(T, V, \mu)$ $= \sum_N \mathbf{e}^{\mu N/k_B T} Z(T, V, N)$	$J(T, V, \mu) = F - \mu N = -pV$ $= -k_B T \ln \Xi(T, V, \mu)$

Table 2: Thermodynamic functions

Thermodynamic functions (Definition)	Natural variables	Total differential
Entropy S	$\langle E \rangle, V, N$	$dS = d\langle E \rangle/T + pdV/T - \mu dN/T$
Internal energy $\langle E \rangle$	S, V, N	$d\langle E \rangle = TdS - pdV + \mu dN$
Enthalpy $H = \langle E \rangle + pV$	S, p, N	$dH = TdS + Vdp + \mu dN$
Helmholtz free energy $F = \langle E \rangle - TS$	T, V, N	$dF = -SdT - pdV + \mu dN$
Gibbs free energy $G = F + pV = N\mu$	T, p, N	$dG = -SdT + Vdp + \mu dN$
Grand potential $J = F - G = -pV$	T, V, μ	$dJ = -SdT - pdV - Nd\mu$

Table 3: Relations of different thermodynamic quantities to the partition function

Canonical partition function (constant number of particles)			
$Z(T, V, N) = \int \Delta\Omega(E, V, N) e^{-\beta E} dE$			
Discrete energy levels: $Z(T, V, N) = \sum_i e^{-\beta E_i}$			
\downarrow			
Free energy			
$F(T, V, N) = -k_B T \ln Z(T, V, N)$		$G(T, P, N) = F(T, V, N) + PV$	
\swarrow Pressure	\searrow Entropy	\swarrow Entropy	\searrow Volume
$P = -\left(\frac{\partial F}{\partial V}\right)_T$	$S = -\left(\frac{\partial F}{\partial T}\right)_V$	$S = -\left(\frac{\partial G}{\partial T}\right)_P$	$V = \left(\frac{\partial G}{\partial P}\right)_T$
\downarrow	\downarrow	\downarrow	\downarrow
Inverse compressibility	Heat capacity (constant V)	Heat capacity (constant P)	Isothermal compressibility
$\frac{1}{\kappa_T} = -V\left(\frac{\partial P}{\partial V}\right)_T$	$C_V = T\left(\frac{\partial S}{\partial T}\right)_V$	$C_P = T\left(\frac{\partial S}{\partial T}\right)_P$	$\kappa_T = -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_T$
$= V\left(\frac{\partial^2 F}{\partial V^2}\right)_T$	$= -T\left(\frac{\partial^2 F}{\partial T^2}\right)_V$	$= -T\left(\frac{\partial^2 G}{\partial T^2}\right)_P$	$= -\frac{1}{V}\left(\frac{\partial^2 G}{\partial P^2}\right)_T$

Chapter 1

Foundations

1.1 Introduction

A macroscopic system is, typically, composed of the order of 10^{23} particles. It is impossible to know, investigate, or describe the exact microscopic behaviour of *each* individual particle in such a system. We must limit our knowledge to the average properties, *thermodynamic quantities* like the temperature or pressure and *correlation functions*.

Statistical mechanics is the bridge between the *microscopic* and *macroscopic* world, it provides methods of calculating the macroscopic properties, like the specific heat, from the microscopic information, like the interaction energy between the particles. The essential ingredient of statistical mechanics is the *probability distribution*, i.e., the collection of occupancies of different *configurations* (microscopic states).

Phase Space. Consider as an example the ideal gas of N particles in a three-dimensional space. The Hamiltonian is

$$\mathcal{H} = \sum_{i=1}^{3N} \frac{p_i^2}{2m} + \sum_{i,j \neq i} V_{ij}; \quad V_{ij} \rightarrow 0. \quad (1.1)$$

(p_i is the component of the momentum, m is the particle mass, and V_{ij} the interaction between particles. The interaction term is necessary for the system of particles to reach thermal equilibrium, but it can be neglected compared to the kinetic energy.). The $3N$ coordinates and $3N$ momenta form the $6N$ dimensional *phase space*. Each point in the phase space represents a *microscopic* state of the

system. The microscopic state changes in time according to the canonical equations of motion:

$$\dot{p}_j = -\frac{\partial \mathcal{H}}{\partial q_j} \quad \text{and} \quad \dot{q}_j = \frac{\partial \mathcal{H}}{\partial p_j}. \quad (1.2)$$

and traces a *trajectory* in the phase space.

Ensembles. The system is completely described if we know the coordinates and momenta of all particles ($6N$ variables - a terrible task). To get an *average* of a physical quantity, say A , one should make the time average over a long segment of the trajectory (a collection of consecutive configurations) of such a system in the phase space,

$$\langle A \rangle = \lim_{\Delta t \rightarrow \infty} \frac{1}{\Delta t} \int_{t_0}^{t_0 + \Delta t} A[\{q_i(t), p_i(t)\}] dt. \quad (1.3)$$

However, usually we know neither the exact microscopic state, i.e., the location of the system in the phase space, nor the trajectory. What we can know, is the macroscopic state (temperature, pressure, volume, ...). Instead of studying a particular microscopic state and the trajectory over the phase space, we investigate an *ensemble of systems*, i.e., a collection of all possible microscopic systems, all belonging to the same macroscopic state.

Ergodicity. To calculate the average of a physical quantity, we substitute the time average over one system by an average over an ensemble of equivalent systems at fixed time:

$$\langle A \rangle = \int_W A(\{q_i, p_i\}) p(\{q_i, p_i\}) d^{3N} q d^{3N} p. \quad (1.4)$$

The integral is over the whole phase space W and p is the probability density in the phase space, it is the probability that a unit volume of the phase space is occupied. So, instead of the *time average*, we make an *ensemble average at fixed time*. When the time and the ensemble averages are equal, we say that the system is *ergodic*. In an alternative but equivalent definition we say that a system is ergodic if it evolves in such a way that it visits with *equal probability* all points of the phase space which are accessible from the initial configuration subject to the constraint of energy conservation. In other words, the trajectory in the phase space of an ergodic system will spend equal time intervals in all regions of the constant energy surface. Most of the physical systems of interest in statistical mechanics are ergodic, so we will in most of the cases replace time averages by the corresponding ensemble averages.

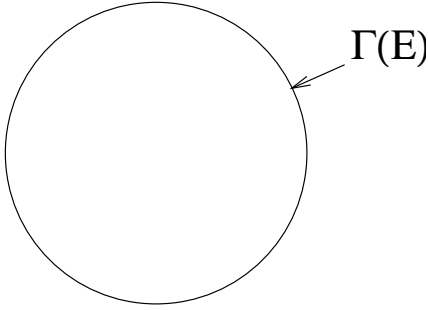


Figure 1.1: Constant-energy surface in the phase space. The area of the surface in the $6N$ -dimensional phase space is $\Gamma(E)$.

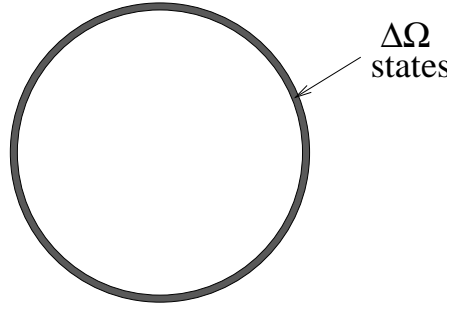


Figure 1.2: Constant-energy shell in the $6N$ -dimensional phase space. The number of states in the shell is $\Delta\Omega(E)$.

For the above example of an isolated ideal gas, the total energy of the system is constant, $\mathcal{H} = E$, the trajectory lies on the constant-energy surface $\Gamma(E)$ in the phase space and the ensemble average of A is:

$$\langle A \rangle = \frac{1}{\Gamma(E)} \int_{\Gamma(E)} A(\{q_i, p_i\}) d^{3N}q d^{3N}p. \quad (1.5)$$

(The integral is over the constant energy surface $\Gamma(E)$.) For quantum systems as well as for classical systems with discrete energy levels, the number of configurations is an extremely irregular function of the energy E . Therefore we introduce a narrow energy interval of the width ΔE and the ergodic hypothesis is generalized so that all the configurations, satisfying the condition that their energy is in the interval between E and $E + \Delta E$, are equally likely to occur. The number of configurations $\Delta\Omega$ with the energy in the interval between E and $E + \Delta E$ is proportional to the volume of the corresponding shell in the phase space. For indistinguishable particles, $\Delta\Omega$ is:

$$\Delta\Omega(E) = \frac{1}{h^{3N} N!} \sum_{E \leq \mathcal{H} \leq E + \Delta E} 1. \quad (1.6)$$

In principle, all systems in Nature obey *quantum statistical mechanics*, the energy levels are discrete, identical particles are indistinguishable and their wave-functions are either symmetric (Bosons) or antisymmetric (Fermions) upon exchange of two particles. *Classical statistical mechanics* is valid only as a special, limiting case when the average occupation of *any* single-particle quantum state is $\ll 1$.

Although we will discuss only statistical mechanics of systems in equilibrium, *i.e.*, *equilibrium statistical mechanics*, there is a growing interest in *nonequilibrium statistical mechanics* which investigates *steady, stationary* states of nonequilibrium systems.

1.2 Ensembles in Statistical Mechanics

An ensemble is a collection of many configurations of a system, all describing the same thermodynamic state. We start with a discussion of isolated systems.

1.2.1 Microcanonical Ensemble and the Entropy (Isolated Systems)

A microcanonical ensemble represents a collection of configurations of isolated systems that have reached thermal equilibrium. A system is isolated from its environment if it exchanges neither particles nor energy with its surrounding. The volume, internal energy and number of particles of such a system are constant and are the same for all configurations that are part of the same microcanonical ensemble. Instead of working with constant energy surface $\Gamma(E)$, we consider all the states with the total energy in a narrow energy interval between E and $E + \Delta E$. All such configurations are lying in a thin shell in the phase space and the total number of configurations is $\Delta\Omega$.

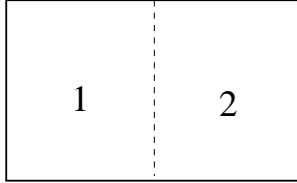
Closely related to the ergodic hypothesis is the *postulate of equal weights* which says that in equilibrium of an isolated system, *all the configurations with equal energy are equally probable*. For a system, belonging to a microcanonical ensemble, therefore, the probability for being in a configuration q_i, p_i with the energy $\mathcal{H}(\{q_i, p_i\})$ is:

$$p(\{q_i, p_i\}) = \begin{cases} 1/\Delta\Omega(E) & E \leq \mathcal{H}(\{q_i, p_i\}) \leq E + \Delta E \\ 0 & \text{otherwise.} \end{cases} \quad (1.7)$$

$\Delta\Omega(E)$ is the number of configurations having the energy in the interval $E \leq \mathcal{H}(\{q_i, p_i\}) \leq E + \Delta E$. All configurations $\{q_i, p_i\}$ which fall into this energy interval are occupied with equal probability $p(\{q_i, p_i\})$.

Entropy and Probability

Probably the most fundamental thermodynamic quantity in statistical physics is the *entropy*. Entropy is a concept, defined only for *macroscopic* systems. A system of only a few particles does not have an entropy. The entropy is a measure of disorder in the system. and is an *additive function of state*.



Let us consider an isolated system composed of two subsystems, which are in contact with each other so that they can exchange energy and particles. As a result of the contact between the two subsystems, an additional uncertainty has been introduced into the system (at the beginning we knew the energy and the number of particles of each individual subsystem, at the end we know only the total energy and number of particles of the whole system).

We *postulate* that the entropy is The entropy is an additive quantity, therefore:

$$S = S_1 + S_2 \quad (1.8)$$

We *postulate* that the entropy is

$$S = k_B \ln \Delta\Omega. \quad (1.9)$$

This relation was discovered by Ludwig Boltzmann and is therefore called the Boltzmann equation. The proportionality constant k_B is the Boltzmann constant and is determined by relating the statistical definition of the entropy to the thermodynamic definition. The Boltzmann equation provides the bridge between thermodynamics (entropy) and statistical mechanics (number of configurations).

The entropy plays the role of thermodynamic potential for microcanonical ensembles, $(-S)$ is minimal for an isolated system in equilibrium (V , E and N are kept constant). From thermodynamics we know that the total differential of the entropy is:

$$TdS(E, V, N) = dE + P dV - \mu dN. \quad (1.10)$$

(This is the *first law of thermodynamics*.) This equation provides a *statistical definition* of the temperature T , pressure P and chemical potential μ :

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E} \right)_{N,V}, \quad (1.11)$$

$$P = T \left(\frac{\partial S}{\partial V} \right)_{N,E}, \quad (1.12)$$

$$\mu = -T \left(\frac{\partial S}{\partial N} \right)_{E,V}. \quad (1.13)$$

T , p and μ are *intensive* quantities, they are independent of N whereas the entropy and the energy E are *extensive* quantities, they are proportional to the number of particles N .

1.2.2 Canonical Ensemble and the Free Energy (Systems at Fixed Temperature)

A *canonical ensemble* is a collection of closed systems at constant temperature. Closed systems have fixed number of particles and volume but can exchange energy with their surroundings. The energy of a closed system is therefore not constant. As an example we can imagine a gas of particles (air molecules, e.g.) in a closed container at constant temperature.

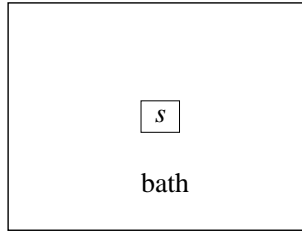


Figure 1.3: A small closed system (s) in contact with a large system, called “bath.” s and bath can exchange energy.

Consider now an isolated system composed of a small closed (sub)system s with N particles in a volume V and of a “heat bath” (b) at temperature T . Let the two subsystems be in contact so that they can exchange energy but no particles. In equilibrium, the system s has equal temperature as the heat bath. Now, the energy of the small system, E_s , is no longer fixed and the system s will be in any energy state E_s , in each one with a different probability p_s . The heat bath must be large ($N_b \gg N_s$), so that the change in the energy of the bath, E_b , during energy exchange between the small system and the bath, is negligible. Evidently, since the compound system (= system + bath) is isolated, the total energy $E_t = E_b + E_s$ is constant and the compound system belongs to a microcanonical ensemble for which the number of configurations in the energy interval between E_t and $E_t + \Delta E$ is:

$$\Delta\Omega_t(E_t) = \Delta\Omega_s(E_s)\Delta\Omega_b(E_t - E_s). \quad (1.14)$$

We first express $\Delta\Omega_b$ with the entropy via the Boltzmann equation (1.9) and then expand $\ln \Delta\Omega_b$ to first order in E_s :

$$k_B \ln \Delta\Omega_b(E_t - E_s) = S_b(E_t - E_s) \approx S_b(E_t) - \left(\frac{\partial S_b}{\partial E} \right) E_s = S_b(E_t) - \frac{E_s}{T}. \quad (1.15)$$

Thus,

$$\Delta\Omega_b(E_t - E_s) = \Delta\Omega_b(E_t) e^{-\beta E_s} = C e^{-\beta E_s}. \quad (1.16)$$

Here, the constant C is independent of E_s and β is the inverse temperature, $\beta = 1/k_B T$.

The (normalized) *probability* for the system s having the energy between E_s and $E_s + dE_s$ is:

$$p(E_s) dE_s = \frac{\Delta\Omega_s(E_s) \Delta\Omega_b(E_t - E_s) dE_s}{\int dE_s \Delta\Omega_s(E_s) \Delta\Omega_b(E_t - E_s)}. \quad (1.17)$$

After expressing $\Delta\Omega_b(E_t - E_s)$ with (1.16) we find the *probability density*:

$$p(E_s) = \frac{1}{Z} \rho(E_s) e^{-\beta E_s}, \quad (1.18)$$

where $\rho(E_s) = \Delta\Omega(E_s)/\Delta E$ is the density of states (number of states per unit energy interval). Z is called the *partition function* and is the normalization constant, depending on the temperature, volume and number of particles of the system s :

$$Z(T, V, N) = \int \rho(E) e^{-\beta E} dE.$$

(1.19)

(Here we omit the index s .) $Z(T, V, N)$ is a Laplace transform of the microcanonical partition function $\Delta\Omega(E) = \rho(E) \Delta E$. For systems with discrete energy levels we replace the integral either by a sum over all the energy levels l :

$$Z(T, V, N) = \sum_l g_l e^{-\beta E_l} \quad (1.20)$$

(g_l is the degeneracy of the energy level l), or by a sum over all configurations i :

$$Z(T, V, N) = \sum_i e^{-\beta E_i}. \quad (1.21)$$

The role of Z is much more important than just being the normalization constant. Z is obtained by integration over the whole phase space. In this way information on microscopic states of the system (coordinates and momenta of individual particles, e.g.) becomes lost but all thermodynamic information on the system (with prescribed temperature, number of particles and volume) is retained and can be obtained from Z , as we shall see below.

The distribution in which the particles are in equilibrium with a heat bath at temperature T is called a canonical distribution similarly as the ensemble defined by this distribution is called a *canonical ensemble*. For a system in equilibrium at a fixed temperature, the system can (at least in principle) access any energy. The probability density $p(E)$ is the same for all states with the same energy but it decreases exponentially with increasing energy.

Free Energy

In this Section we will see the relation between the average of a physical quantity of a canonical ensemble and the partition function Z .

The average energy $\langle E \rangle$ and average generalized force $\langle X \rangle$ in a canonical distribution are given by

$$\langle E \rangle = \frac{1}{Z} \int E \rho(E) e^{-\beta E} dE = -\frac{\partial}{\partial \beta} \ln Z, \quad (1.22)$$

$$\langle X \rangle = \frac{1}{Z} \int X \rho(E) e^{-\beta E} dE = \int \frac{\partial E}{\partial x} e^{-\beta E} dE = -\frac{\partial}{\partial x} \frac{1}{\beta} \ln Z. \quad (1.23)$$

x is the generalized coordinate, conjugate to the generalized force. Examples of conjugate forces and coordinates are: pressure vs. volume in gases, magnetic field vs. magnetization in magnetic systems; the generalized force is intensive while the coordinate is an extensive quantity.

We introduce the (Helmholtz) *free energy* as:

$$F(T, V, N) = -k_B T \ln Z(T, V, N). \quad (1.24)$$

Like the Boltzmann equation (1.9) in case of the microcanonical ensemble, also this equation relates thermodynamics (free energy) and statistical mechanics (partition function).

Relation between F and S : We take the well-known relation from thermodynamics:

$$F(T, V, N) = \langle E(S, V, N) \rangle - TS, \quad (1.25)$$

$F(T, V, N)$ is a Legendre transform of $\langle E(S, V, N) \rangle$, by which the independent variables are changed from (S, V, N) to (T, V, N) .

The total differential of F is:

$$dF = -SdT - PdV + \mu dN. \quad (1.26)$$

Another Legendre transform relates the Helmholtz and the Gibbs free energies $F(T, V, N)$ and $G(T, P, N)$:

$$G(T, P, N) = F(T, V, N) + PV \quad (1.27)$$

For a list of thermodynamic functions and their derivatives, see Tables 1 to 3 in the Appendix B. Exercise: Construct these tables for magnetic systems!

Stability and convexity of the free energies (See Fig. 1.4). The Helmholtz free energy $F(T, V)$ (for simplicity, we will omit N) is a concave (i.e, the second derivative is negative) function of T and convex function of V and the Gibbs free energy $G(T, P)$ is always a concave function of both T and P .

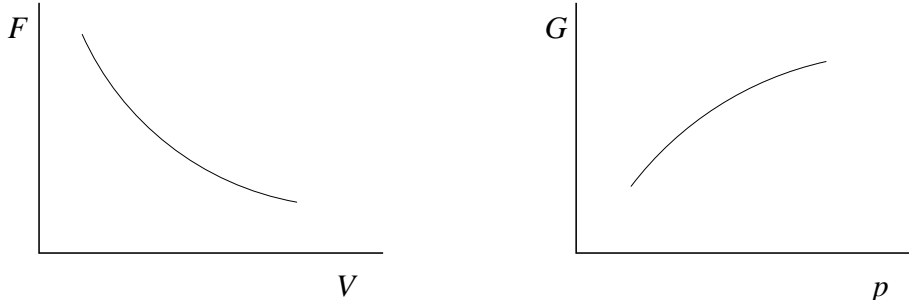


Figure 1.4: Schematic behaviour of the Helmholtz and Gibbs free energies.

These properties follow from the fact that the system must be thermally and mechanically stable, therefore the specific heat and the compressibility must be positive quantities. We begin by noting that

$$S = -\left(\frac{\partial G}{\partial T}\right)_p = -\left(\frac{\partial F}{\partial T}\right)_V. \quad (1.28)$$

It follows that

$$\left(\frac{\partial^2 G}{\partial T^2}\right)_p = -\left(\frac{\partial S}{\partial T}\right)_p \equiv -\frac{1}{T}C_p \leq 0 \quad (1.29)$$

and

$$\left(\frac{\partial^2 F}{\partial T^2}\right)_V = -\left(\frac{\partial S}{\partial T}\right)_V \equiv -\frac{1}{T}C_V \leq 0. \quad (1.30)$$

Hence, $G(T, p)$ and $F(T, V)$ are concave functions of T .

Together with equation (1.26) we see that for a system in equilibrium at constant T, V, N the distribution of a system over possible configurations is such that F is *minimal*. Therefore F is also called canonical or Gibbs *potential*.

The next two properties are obtained by differentiating $G(T, P)$ and $F(T, V)$ with respect to pressure and volume, respectively:

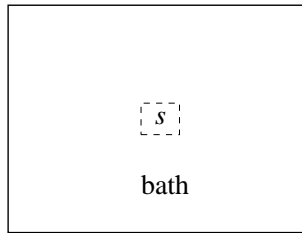
$$\left(\frac{\partial^2 G}{\partial P^2}\right)_T = \left(\frac{\partial V}{\partial P}\right)_T \equiv -V\kappa_T \leq 0 \quad (1.31)$$

$$\left(\frac{\partial^2 F}{\partial V^2}\right)_T = -\left(\frac{\partial P}{\partial V}\right)_T \equiv \frac{1}{V\kappa_T} \geq 0. \quad (1.32)$$

Hence, $G(T, P)$ is a concave function of pressure P and $F(T, V)$ is a convex function of volume V .

The above properties are valid also for systems that undergo phase transitions.

1.2.3 Grand Canonical Ensemble (Open systems)



Consider now a more general case with the system of volume V_s being not only in contact with a heat bath at temperature T , but also exchanging particles with the bath. As an example we can imagine a room at constant temperature with an open window. In this case, neither the energy nor the number of particles in the system (air molecules in the room, e.g.) are fixed. Such a system is described by the *grand canonical distribution*.

We proceed like in case of the canonical ensemble, we consider the total system being part of a microcanonical ensemble, therefore the number of configurations of the total system is:

$$\Delta\Omega(E_t, N_t) = \Delta\Omega_s(E_s, N_s) \Delta\Omega_b(E_t - E_s, N_t - N_s). \quad (1.33)$$

We express $\Delta\Omega_b$ with the entropy via the Boltzmann equation and then expand $\ln \Delta\Omega_b$ to first order in E_s and N_s :

$$\begin{aligned} k_B \ln \Delta\Omega_b(E_t - E_s, N_t - N_s) &= S_b(E_t - E_s, N_t - N_s) \\ &\approx S_b(E_t, N_t) - \left(\frac{\partial S_b}{\partial E} \right) E_s - \left(\frac{\partial S_b}{\partial N} \right) N_s \\ &= S_b(E_t, N_t) - \frac{E_s}{T} + \frac{\mu N_s}{T}. \end{aligned}$$

Thus,

$$\Delta\Omega_b(E_t - E_s, N_t - N_s) = C e^{-\beta(E_s - \mu N_s)}. \quad (1.34)$$

Eventually we find for the probability density of a system having the energy E and N particles:

$$p(E, N) = \frac{1}{\Xi} \rho(E, N) e^{-\beta(E - \mu N)}, \quad (1.35)$$

where the normalization constant Ξ is called the *grand (canonical) partition function*:

$$\Xi(T, V, \mu) = \sum_N \int \rho(E, N) e^{-\beta(E - \mu N)} dE. \quad (1.36)$$

(Again, we omit the index s .) The grand partition function Ξ is related to the partition function Z :

$$\Xi(T, V, \mu) = \sum_N Z(T, V, N) e^{\beta\mu N}. \quad (1.37)$$

Thermodynamic quantities in the grand canonical ensemble

In analogy to the free energy in the case of a canonical ensemble, we introduce the grand (canonical) potential J which will be used for open systems:

$$J(T, V, \mu) = -k_B T \ln \Xi \quad (1.38)$$

The total differential of J is:

$$dJ = -SdT - pdV - Nd\mu. \quad (1.39)$$

The entropy can now be expressed as

$$S = - \left(\frac{\partial J}{\partial T} \right)_{V, \mu} = k_B \ln \Xi + k_B T \left(\frac{\partial \ln \Xi}{\partial T} \right)_{V, \mu}, \quad (1.40)$$

the pressure is:

$$p = -\left(\frac{\partial J}{\partial V}\right)_{T,\mu} = k_B T \left(\frac{\partial \ln \Xi}{\partial V}\right)_{T,\mu}, \quad (1.41)$$

and the number of particles is

$$N = -\left(\frac{\partial J}{\partial \mu}\right)_{T,V} = k_B T \left(\frac{\partial \ln \Xi}{\partial \mu}\right)_{T,V}. \quad (1.42)$$

For a system in equilibrium at constant T , V , and μ , the distribution of a system over possible configurations is such that the potential J is minimal.

1.3 Analogy Between Fluids and Magnetic Systems

Above, we have mainly discussed fluids (liquids, gases). When we want to consider, e.g., magnetic systems, the above relations between thermodynamic quantities are easily written down for them as well. The potential energy of a magnetic dipole μ in a magnetic field H is

$$W = -\mu H \quad (1.43)$$

and the work required to increase the magnetization of a system by dM is:

$$-H d\mu. \quad (1.44)$$

Hence, the first law of thermodynamics for magnets reads:

$$T dS = dE - H dM \quad (1.45)$$

Comparing this equation with 1.10 we see that we only have to replace the "generalized force" - pressure P by the "force", acting on the magnetic moment, i.e., by the negative magnetic field:

$p \rightarrow -H.$

(1.46)

Similarly, the "generalized coordinate" V is replaced by the "generalized coordinate" of magnetic systems, M :

$V \rightarrow M.$

(1.47)

Notice that P and H are intensive quantities whereas V and M are extensive.

1.4 Fluctuations

The microcanonical ensemble describes a collection of isolated systems where neither energy nor the number of particles can fluctuate.

Energy fluctuations in the canonical ensemble: The canonical ensemble describes a system in contact with a heat bath. For such a system, the energy is not exactly fixed, it can fluctuate around $\langle E \rangle$. We will now evaluate the square of the energy fluctuations,

$$(\Delta E)^2 = \langle (E - \langle E \rangle)^2 \rangle = \langle E^2 \rangle - \langle E \rangle^2. \quad (1.48)$$

In the canonical ensemble, the thermal average of the energy is

$$\langle E \rangle = -\frac{1}{Z} \frac{\partial Z}{\partial \beta} \quad (1.49)$$

and the thermal average of the energy squared is:

$$\langle E^2 \rangle = \frac{1}{Z} \int E^2 \rho(E) e^{-\beta E} dE = \frac{1}{Z} \frac{\partial^2 Z}{\partial \beta^2}. \quad (1.50)$$

So, the square of the energy fluctuation is

$$(\Delta E)^2 = \frac{1}{Z} \frac{\partial^2 Z}{\partial \beta^2} - \left(\frac{1}{Z} \frac{\partial Z}{\partial \beta} \right)^2 = \frac{\partial^2 \ln Z}{\partial \beta^2}. \quad (1.51)$$

After expressing this with $\langle E \rangle$ (see paragraph 2.3.3) we get:

$$(\Delta E)^2 = -\frac{\partial \langle E \rangle}{\partial \beta} = k_B T^2 \frac{\partial \langle E \rangle}{\partial T}. \quad (1.52)$$

$$(\Delta E)^2 = k_B T^2 C_V. \quad (1.53)$$

Here, C_V is the *heat capacity*, it is an extensive quantity and is proportional to N . The relative fluctuation of the energy is thus:

$$\frac{\Delta E}{\langle E \rangle} = \sqrt{\frac{k_B c_V}{N}} \frac{T}{\varepsilon}, \quad (1.54)$$

where ε is the energy per particle and c_V is the specific heat at constant volume, $c_V = C_V/N$. The relative energy fluctuation is $\propto 1/\sqrt{N}$ and it vanishes as $N \rightarrow \infty$, see Fig. 1.5. For macroscopic systems (i.e., in the limit $N \rightarrow \infty$), the

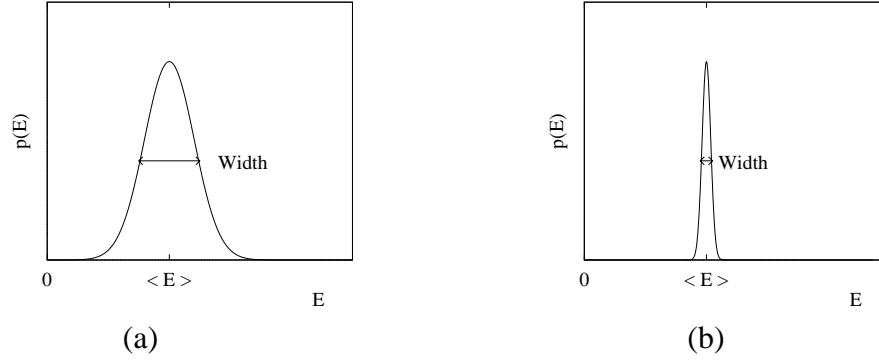


Figure 1.5: Fluctuation in the energy of a small (a) and of a large (b) system.

distribution of energies in a canonical ensemble is so sharply peaked around the average energy that the *canonical ensemble is equivalent to the microcanonical ensemble*. For later use we write Eq. (1.54) in the form:

$$c_v = \frac{\beta}{NT} \left(\langle E^2 \rangle - \langle E \rangle^2 \right) \quad (1.55)$$

Fluctuations in the grand canonical ensemble: Now the number of particles is not fixed, so we shall calculate ΔN (and ΔE). The calculation is analogous to the above calculation for the canonical ensemble. We start with the grand partition function

$$\Xi(T, V, \mu) = \sum_N \int \rho(E) e^{-\beta(E - \mu N)} dE \quad (1.56)$$

and differentiate it twice with respect to μ by keeping β and V constant,

$$\langle N \rangle = \frac{1}{\beta \Xi} \frac{\partial \Xi}{\partial (\mu)}, \quad (1.57)$$

$$\langle N^2 \rangle = \frac{1}{\beta^2 \Xi} \frac{\partial^2 \Xi}{\partial \mu^2}. \quad (1.58)$$

The square of the particle number fluctuation is:

$$(\Delta N)^2 = \langle N^2 \rangle - \langle N \rangle^2 = \left(\frac{\partial^2 \ln \Xi}{\beta^2 \partial \mu^2} \right)_{T,V}. \quad (1.59)$$

For fluids at constant T and V , we obtain from the thermodynamic relation (See Table 1) $pV = k_B T \ln \Xi$:

$$\frac{\partial^2 \ln \Xi}{\partial \mu^2} = \frac{V}{k_B T} \frac{\partial^2 p}{\partial \mu^2}. \quad (1.60)$$

We assume that the (Helmholtz) free energy can be written as $F(T, V, N) = N f(T, v)$, where $f(T, v)$ is the free energy per particle and depends on T and $v \equiv V/N$. Then,

$$\mu(T, p) = f(T, v) + pv \quad (1.61)$$

and

$$\frac{\partial^2 p}{\partial \mu^2} = \frac{\partial}{\partial \mu} \left(\frac{1}{v} \right) = -\frac{1}{v^2} \frac{\partial v}{\partial \mu}, \quad (1.62)$$

where

$$\frac{\partial \mu}{\partial v} = \frac{\partial \mu}{\partial p} \frac{\partial p}{\partial v} = v \frac{\partial p}{\partial v} \equiv -\frac{1}{\kappa_T}. \quad (1.63)$$

κ_T is the isothermal compressibility. We get finally that

$$\langle N^2 \rangle - \langle N \rangle^2 = \frac{\langle N \rangle k_B T \kappa_T}{v}. \quad (1.64)$$

The square of the fluctuation in the number of particles is proportional to the (isothermal) compressibility! This equation shows also that $\Delta N \propto \sqrt{N}$ and that $\Delta N / \langle N \rangle$ vanishes as $N \rightarrow \infty$. We conclude that the grand canonical ensemble is equivalent to the canonical ensemble when $\langle N \rangle \rightarrow \infty$. The decision about which ensemble is to be used to describe the investigated system does not depend on whether the system is closed or open, but it is a matter of convenience. For later use we rewrite Eq. 1.64 in a different form:

$$\kappa_T = \frac{v\beta}{\langle N \rangle} \left(\langle N^2 \rangle - \langle N \rangle^2 \right) \quad (1.65)$$

The limit $N \rightarrow \infty$, in which the relative fluctuations become negligible, is called the *thermodynamic limit*. So far we assumed c_V or κ_T to be constant. At a continuous phase transition however (see Section 1.5), these quantities and consequently also the corresponding fluctuations diverge on *all lengthscales*. This makes life of experimentalists and of those who make simulations on computers (both can deal with finite systems only) very difficult.

The fluctuations of N around $\langle N \rangle$ or E around $\langle E \rangle$ are - of course - time-dependent processes! Therefore we have to assume that the systems are ergodic in order to identify the *time average* with an *ensemble average* which we actually calculated.

Fluctuation-Dissipation Theorem

Above, we have shown that $(\Delta E)^2$ is proportional to the specific heat c_V and that $(\Delta N)^2$ is proportional to the compressibility. In the problem 3 we also see that $(\Delta M)^2$ is proportional to the susceptibility. All these examples indicate that there is a general relation between the *fluctuations* of the generalized coordinate and the *response* of this coordinate to the perturbation of the conjugate generalized force.

This is the message of the *fluctuation-dissipation theorem* which in a most general way relates the fluctuations of a physical quantity of a system in equilibrium to a dissipation process which is realized when the system is subject to an external force driving it away from equilibrium. The theorem makes it possible to find the *non-equilibrium* properties of a system from the analysis of thermal fluctuations of the system in *equilibrium*. It is a bridge between non-equilibrium and equilibrium statistical mechanics.

Now let us have a look at magnetic systems. The magnetic susceptibility, i.e., the response of the magnetization M to the external magnetic field H is:

$$\chi = \frac{\beta}{N} (\langle M^2 \rangle - \langle M \rangle^2). \quad (1.66)$$

If we write $M = \sum_i S_i$, the susceptibility becomes:

$$\chi = \frac{\beta}{N} \sum_{i,j} [\langle S_i S_j \rangle - \langle S_i \rangle \langle S_j \rangle]. \quad (1.67)$$

The summand on the right-hand side is the spin-spin *correlation function*,

$$\Gamma_{ij} = [\langle S_i S_j \rangle - \langle S_i \rangle \langle S_j \rangle]. \quad (1.68)$$

which describes the correlations in the spatial fluctuations of the spin around its thermal average. Because of the interaction between the nearest-neighbouring spins, one expects that the spins are *correlated* on short distances and that their fluctuations are less and less correlated as the distance increases. The correlation function thus decays with the distance. Often Γ will be assumed to have the form

$$\Gamma(r) \propto r^{-p} e^{-r/\xi}, \quad (1.69)$$

where the decay length ξ is called the *correlation length*. As we shall see later, ξ is a temperature-dependent quantity which plays a crucial role in phase transitions, it diverges at a continuous phase transition.

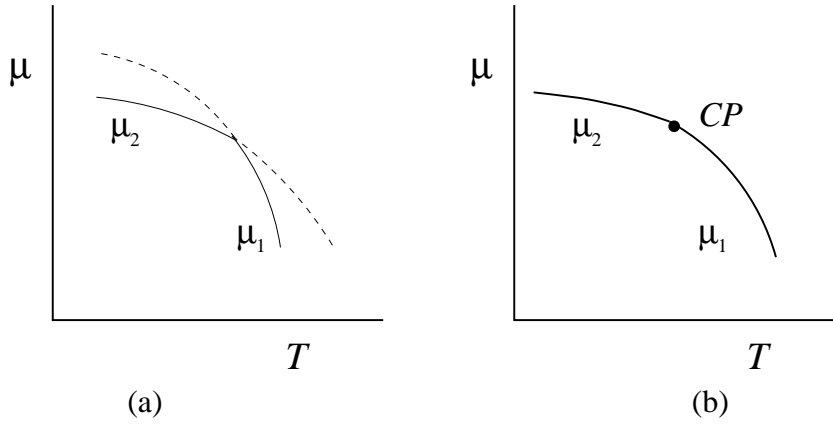


Figure 1.6: Chemical potentials of two phases in contact. When $\mu_1 = \mu_2$, the two phases coexist. In figure (a) the chemical potentials cross and the transition is first order. Dashed portions represent unstable states. In figure (b) the transition is continuous and the two phases meet in the critical point CP .

1.5 Phase Transitions

So far we have considered only homogeneous systems, i.e., single phases. In Nature there are, however, many systems that exist in different phases. Examples are solid-liquid-gas, magnetic, ferroelectric, \dots , systems. In this Section we will discuss transitions between such phases.

When two phases, say 1 and 2 are in contact with each other, there is an *interface* separating the phases. In equilibrium, the chemical potentials of both phases in contact are equal,

$$\mu_1(T, p_1, \dots) = \mu_2(T, p_2, \dots) \quad (1.70)$$

and the two phases *coexist* (See Figure 1.6).

Classification of Phase Transitions

As one of the parameters (e.g., the temperature) is varied in such a way that the coexistence line is crossed, the system undergoes a *phase transition* between the two phases. At the phase transition a derivative of the free energy (or some other

thermodynamic potential) is discontinuous. The free energy (thermodynamic potential) becomes non-analytic in this point. The transition can be either *first order* or *continuous*.

(a) *First-order phase transitions*. The conditions for first-order phase transitions are that the chemical potentials are equal,

$$\Delta\mu = \mu_1(T, p, \dots) - \mu_2(T, p, \dots) = 0, \quad (1.71)$$

and that their derivatives are different,

$$\frac{\partial\Delta\mu}{\partial T} \neq 0 \quad \text{and} \quad \frac{\partial\Delta\mu}{\partial p} \neq 0, \quad (1.72)$$

so that the chemical potentials or the free energies *cross* at the transition. The phase with the lower chemical potential is *stable* and the other phase is *metastable*. An example of a first-order phase transition is the liquid-gas transition.

(b) *Continuous phase transitions*. A transition is continuous when the chemical potentials of both phases are equal and when their first derivatives are also equal,

$$\Delta\mu = \mu_1(T, p, \dots) - \mu_2(T, p, \dots) = 0, \quad (1.73)$$

$$\frac{\partial\Delta\mu}{\partial T} = 0 \quad \text{and} \quad \frac{\partial\Delta\mu}{\partial p} = 0, \quad (1.74)$$

so that the chemical potentials or the free energies have a common tangent at the transition. There is no latent heat in this type of transition.

Instead of continuous transitions, some authors write about *second*, *third*, and higher order transitions, depending on which derivative of the chemical potential diverges or becomes discontinuous. This classification is somewhat arbitrary, therefore we will distinguish only between the continuous and first-order transitions.

1.5.1 Order Parameter

As an example, let us consider a simple ferromagnetic material. At high temperatures, there is no magnetization and the system is rotationally invariant. At low

temperatures, when *spontaneous magnetization* occurs, the magnetization direction defines a preferred direction in space, destroying the rotational invariance. We call this *spontaneous symmetry breaking* and it takes place at the *critical temperature* (for ferromagnets it is the *Curie point*). Since a symmetry is either absent or present, the two phases must be described by different functions of thermodynamic variables, which cannot be continued analytically across the critical point.

Because of the reduction in symmetry, an additional parameter is needed to describe the thermodynamics of the low-temperature phase. This parameter is called the *order parameter*. It is usually an extensive thermodynamic variable. In the case of ferromagnets, the order parameter is the thermal average of the (total) magnetization, $\langle M \rangle$. For the gas-liquid transition, the order parameter is either the volume or the density difference between the two phases. Notice that the density is not an extensive quantity! Notice also that there is no obvious symmetry change in the gas-liquid transition.

The basic idea of phase transitions is that near the critical point, the order parameter is the only relevant thermodynamic quantity.

Some examples of the order parameters and conjugate fields are given in the following table:

Systems	Order Parameter	Conjugate Field
Gas-liquid	$V_G - V_L$	$p - p_c$
Ferromagnets	M	H
Antiferromagnets	$M_{staggered}$	$H_{staggered}$
Superfluids	$\langle \psi \rangle$ (Condensate wave fn.)	Not physical
Superconductors	Δ (Gap parameter)	Not physical

1.5.2 Critical Point

1. *Liquid-gas transitions.* In Figure 1.7(a) a typical phase diagram of a solid-liquid-gas system in the $P-T$ plane at constant V and N is shown. Here we would like to concentrate on the liquid-gas transition. The liquid and the gas phases coexist along a line where the transition between these two phases is first order.

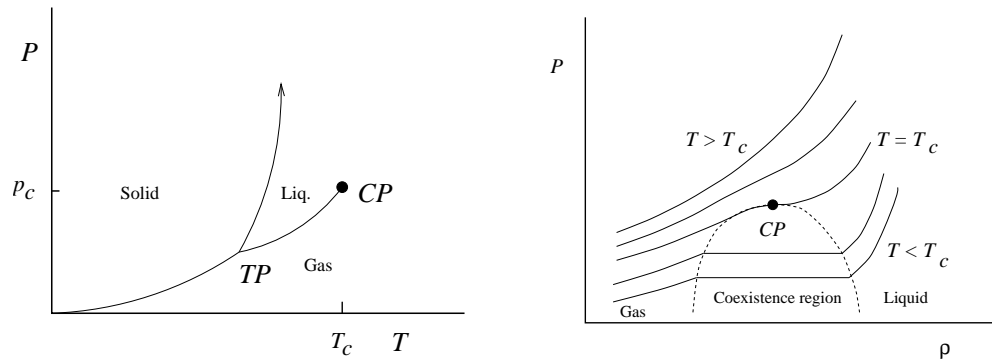


Figure 1.7: Phase diagram of a fluid. CP is the critical point; TP is the triple point.

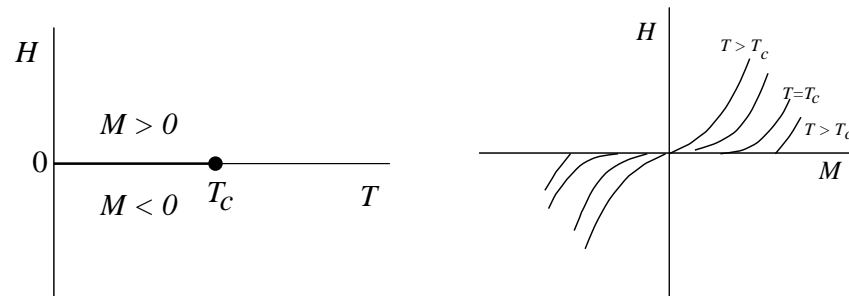


Figure 1.8: Phase diagram of a ferromagnet.

If T is increased, the coexistence line terminates in the *critical point*, where the transition between the two phases is continuous. Beyond the critical point, there is no transition between the two phases, there is no singularity in any physical quantity upon going from one phase to the other. It is also instructive to consider the isotherms in the $P - \rho$ plane at constant T and N , see Fig. 1.7(b). At low temperature, there is a large difference between the gas and liquid densities, ρ_G and ρ_L , but as the critical temperature is approached, this difference tends to zero. The existence of a quantity that is non-zero below the critical temperature and zero above it, is a common feature of critical points in a variety of different physical systems. $\rho_L - \rho_G$ is the *order parameter* for the liquid-gas critical point.

2. *Ferromagnets.* A typical phase diagram of a ferromagnet is shown in Fig. 1.8(a). The order parameter is the magnetization M , i.e., the total magnetic moment. At low temperature, M is ordered either "up" or "down", depending on the external field H . When H changes sign, the system undergoes a first-order phase transition (bold line) in which the magnetization changes its direction discontinuously [Fig. 1.8(b)]. As T is increased, the *critical point* is reached, at which M changes continuously upon variation of H but its derivative with respect to H , i.e., the *susceptibility diverges*. At still higher T , there is no phase transition, the spins are simply disordered, this is the paramagnetic region.

1.5.3 Critical Exponents

In this Section the behaviour of the thermodynamic quantities *near* the critical point is discussed. We suppose that very near the critical point, any thermodynamic quantity can be decomposed into a "regular" part, which remains finite (not necessarily continuous) and a "singular part" that can be divergent or have divergent derivatives. The singular part is assumed to be proportional to some power of $|T - T_C|$ (or $p - p_c$).

Definition of the critical exponents. For clarity, we will define the critical exponents on a ferromagnet. However, their definition will be general and will be valid for arbitrary system that undergoes a continuous phase transition. The most important critical exponents are:

β = *Order parameter exponent.* When $H = 0$, the magnetization M in a ferromagnet is a decreasing function of T and vanishes at T_C (See Fig. 1.9).

For T very close to (but below) T_C , M can be written as

$$M \sim (T_C - T)^\beta. \quad (1.75)$$

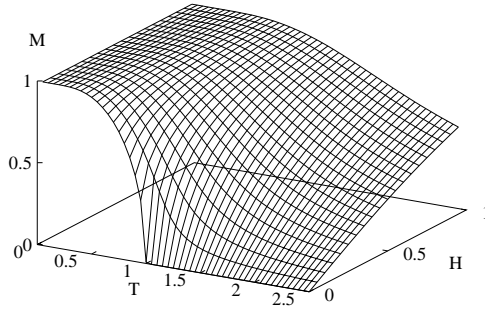


Figure 1.9: Order parameter vs. temperature and field.

β is the *order parameter exponent* which describes the behaviour of M near T_C . Here and later, \sim means "the singular part is proportional to".

$\delta = \text{Exponent of the critical isotherm}$. When $T = T_C$, M vanishes for $H = 0$ but increases very rapidly if H increases. For small H , the singular part of M is proportional to:

$$M \sim H^{1/\delta}, \quad (1.76)$$

where $\delta > 1$.

$\gamma = \text{Susceptibility exponent}$. As T_C is approached either from above or from below, the susceptibility diverges. The divergence is described with the exponent γ :

$$\chi \sim |T - T_C|^{-\gamma}. \quad (1.77)$$

Notice that the exponent γ is the same for $T > T_C$ and $T < T_C$ whereas the proportionality constants are different.

$\alpha = \text{Specific heat exponent}$. The specific heat at constant volume (for ferromagnetic systems one has to take $H = 0$) has a singularity at T_C . This singularity is described by the exponent α :

$$C_V \sim |T - T_C|^{-\alpha}. \quad (1.78)$$

As usually, the exponent α is the same for $T > T_C$ and $T < T_C$ whereas the proportionality constants are different.

$\nu = \text{Correlation length exponent}$. As already mentioned, the correlation length ξ is a temperature- dependent quantity. As we shall see later, ξ diverges at the

critical point with the exponent ν :

$$\xi \sim |T - T_C|^{-\nu}. \quad (1.79)$$

$\eta =$ *Correlation function exponent at the critical isotherm*. At T_C , ξ diverges and the correlation function Γ decays as a power law of r ,

$$\Gamma \sim r^{d-2+\eta}, \quad (1.80)$$

where d is the dimensionality of the system.

Scaling. As the name suggests, the scaling has something to do with the change of various quantities under a change of length scale. Important scaling laws concerning the thermodynamic functions can be derived from the simple (but strong) assumption that, near the critical point, the correlation length ξ is the only characteristic length of the system, in terms of which all other lengths must be measured. This is the "scaling hypothesis."

With the aid of the scaling hypothesis we are able to get relations between the critical exponents. The most common scaling relations are:

$$\begin{aligned} \gamma &= \beta(\delta - 1) \\ \gamma &= \nu(2 - \eta) \\ \alpha + 2\beta + \gamma &= 2 \\ \nu d &= 2 - \alpha \end{aligned} \quad (1.81)$$

so that only two of the above exponents are independent. The last four equations are called the *scaling laws*. We shall prove some of the scaling laws later.

Universality. Some quantities depend on the microscopic details of the system. Examples of such quantities are:

- The critical temperature depends on the strength of interaction.
- The proportionality constants (e.g., for the temperature dependence of the magnetization) also depend on the microscopic properties of the system.

On the other hand, the *critical exponents* and the functional form of the equation of state, of the correlation functions, etc. are independent of many details of the system. They depend only on global features, such as:

- The spatial dimensionality,
- symmetry (isotropic, uniaxial, planar,...),

Systems that have the same critical exponents and the same functional form of the equation of state, etc, belong to the same *universality class*. For example,

the Ising model in $d = 2$ forms one universality class, whereas in $d = 3$ it forms another universality class, the Ising models in $d = 2$ and $d = 3$ have different critical exponents because they have different spatial dimensionality. However, on a first glance very different physical systems often belong to the same universality class. Example: binary alloys and some magnetic systems can be described with mathematically similar models and they belong to the same universality class. Roughening of some crystal surfaces at high temperature (but below the melting temperature) and the $d = 2$ Coulomb gas are members of another universality class.

Chapter 2

Classical Models

In this Chapter we will discuss models that can be described by classical statistical mechanics. We will concentrate on the classical spin models which are used not only to study magnetism but are valid also for other systems like binary alloys or lattice gases.

In principle, all particles obey quantum statistical mechanics. Only when the temperature is so high and the density is so low that the average separation between the particles is much larger than the thermal wavelength λ ,

$$\left(\frac{V}{N}\right)^{1/3} \gg \lambda, \text{ where } \lambda = \left(\frac{h^2}{2\pi m k_B T}\right)^{1/2}, \quad (2.1)$$

the quantum statistical mechanics reduces to classical statistical mechanics and we can use a classical, Maxwell-Boltzmann distribution function (instead of Fermi–Dirac or Bose–Einstein).

For magnetic materials, the situation is somewhat different. The Hamiltonian of a magnetic system is a function of spin operators which can usually not be directly approximated by classical vectors. Quantum models in d dimensions can be mapped onto classical models in $d' = (d + 1)$ dimensions. The classical limit of the Heisenberg model, however, can be constructed for large eigenvalues of the spin operator by replacing the spin operators by three-dimensional classical vectors.

There is another quantum-mechanical effect we must discuss. In quantum mechanics identical particles are indistinguishable from each other whereas in classical mechanics they are distinguishable. Therefore, the partition function of a system of non-interacting, non-localized particles (ideal gas!) is not just the product of single-particle partition functions, as one would expect from classical statistical

mechanics, but we must take into account that the particles are indistinguishable. We must divide the total partition function by the number of permutations between N identical particles, $N!$,

$$Z_N = \frac{1}{N!} (Z_1)^N, \quad (2.2)$$

where Z_1 is the one-particle partition function. The factor $N!$ is a purely quantum effect and could not be obtained from classical statistical mechanics.

In (insulating) magnetic systems we deal with *localized* particles (spins) and permutations among particles are not possible. Therefore we must not divide the partition function by $(N!)$.

A Hamiltonian of a simple magnetic system is:

$$\mathcal{H} = -\frac{1}{2} \sum_{i,j} J_{i,j} \hat{\vec{S}}_i \hat{\vec{S}}_j - \vec{H} \sum_i \hat{\vec{S}}_i. \quad (2.3)$$

The first term is the interaction energy between the spins, $J_{i,j}$ being the (exchange) interaction energy between the spins at the sites i and j . The sum is over all *bonds* between the spins. $\hat{\vec{S}}$ is the spin operator; in general it is a vector. The *order parameter* is equal to the thermal and quantum-mechanical average of $\hat{\vec{S}}$, $\vec{m} = \langle \langle \hat{\vec{S}} \rangle \rangle$. The thermal average of the first term in Eq. (2.3) is $E(\mathcal{S}, \vec{m}) = \langle E(\mathcal{S}, \{S\}) \rangle$. (Here, \mathcal{S} is the entropy and $\{S\}$ is a spin configuration. The thermal average has to be taken over all spin configurations.) $E(\mathcal{S}, \vec{m})$ is the *internal energy* of a magnetic system.

The second term in (2.3) is the interaction with external magnetic field \vec{H} . The thermal average of this term corresponds to PV in fluids. We see that $\langle \mathcal{H}(\mathcal{S}, \vec{H}) \rangle$ is a Legendre transformation of $E(\mathcal{S}, \vec{m})$ and is thus equivalent to the enthalpy of fluids.

2.1 Real Gases

In the ideal gas we neglected the collisions between the particles. In reality, the atoms (molecules) have a finite diameter and when they come close together, they *interact*. The Hamiltonian of interacting particles in a box is:

$$\mathcal{H} = \sum_i \left[\frac{\vec{p}_i^2}{2m} + U(\vec{r}_i) \right] + \sum_{i>j} u(r_i - r_j), \quad (2.4)$$

$u(r_i - r_j) = u(r)$ is the interaction energy and U the potential of the box walls. The particles are polarizable and this leads to an attractive van der Waals interaction between the induced electric dipole moments. Such an interaction energy is proportional to r^{-6} , where r is the interatomic distance. At short distances, however, the electrons repel because of the Pauli exclusion principle. Although the repulsive potential energy usually decays exponentially with r , we will write it in the form:

$$u(r) = \begin{cases} \infty & r < 2r_0 \\ -\epsilon(2r_0/r)^6 & r > 2r_0 \end{cases} \quad (2.5)$$

($2r_0$ is the minimal distance to which the particles of "hard-core" radius r_0 can approach.)

The partition function of N interacting identical particles is

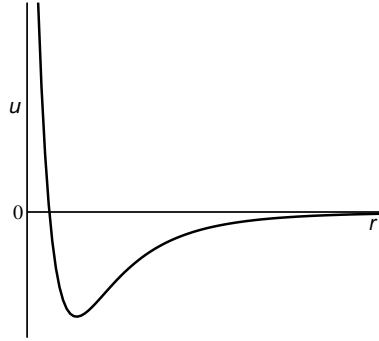


Figure 2.1: Interatomic potential energy $u(r)$ in real gases.

$$Z(T, V, N) = \frac{1}{h^{3N} N!} \int \cdots \int d^3 p_1 \cdots d^3 p_N \int \cdots \int d^3 r_1 \cdots d^3 r_N \times e^{-\beta [\sum_i (p_i^2/2m + U) + \sum_{i>j} u(r_{i,j})]}. \quad (2.6)$$

After integrating over all momenta, the partition function becomes:

$$Z(T, V, N) = \frac{1}{N!} \left(\frac{2\pi m k_B T}{h^2} \right)^{3N/2} Z_I(T, V, N), \quad (2.7)$$

where Z_I carries all the interactions between the particles. We calculate the partition function in a mean-field approximation in which all the correlations are neglected. The potential acting on one particle is calculated by assuming a uniform

spatial distribution of other particles,

$$Z_I = \left[\int_0^{V_1} d^3 r_1 e^{-\beta \Phi} \right]^N \quad (2.8)$$

V_1 is the volume, available to one particle, $V_1 = V - (N/2)(4\pi/3)(2r_0)^3 = V - Nb$. Here we subtracted from V the volume occupied by other particles. $N/2$ is the number of pairs of particles. The mean-field potential Φ is:

$$\Phi = - \int_{2r_0}^{\infty} \epsilon (2r_0/r)^6 (N/V) 4\pi r^2 dr = - \frac{2Na}{V} \quad (2.9)$$

The resulting partition function is

$$Z(T, V, N) = \frac{1}{N!} \left(\frac{2\pi m k_B T}{h^2} \right)^{3N/2} (V - Nb)^N e^{\beta N^2 a/V} \quad (2.10)$$

($a = (2\pi/3)\epsilon (2r_0)^3$, $b = (2\pi/3)(2r_0)^3$), the Helmholtz free energy is:

$$F(T, V, N) = -k_B T \ln Z = -Nk_B T \ln \left[\frac{V - Nb}{N} \left(\frac{2\pi m k_B T}{h^2} \right)^{3/2} \right] - \frac{N^2 a}{V}, \quad (2.11)$$

and the equation of state is

$$p(T, V, N) = - \left(\frac{\partial F}{\partial V} \right)_{T,N} = \frac{Nk_B T}{V - Nb} - \frac{N^2 a}{V^2}. \quad (2.12)$$

The last equation can be written in a more familiar form:

$$\boxed{\left(p + \frac{N^2 a}{V^2} \right) (V - Nb) = Nk_B T.} \quad (2.13)$$

This is the familiar *van der Waals equation*. We derived the van der Waals equation of state in a mean-field approximation, we neglected all the correlations between the molecules. Nevertheless, this equation qualitatively correctly describes the liquid-gas transition and it also predicts the existence of the critical point. However, as we neglected the correlations, which are essential in the vicinity of the critical point, it does not predict the correct critical behaviour.

The critical point is located at the inflexion point of the isotherm, $(\partial p / \partial V)_T = (\partial^2 p / \partial V^2)_T = 0$, and is located at:

$$\begin{aligned} V_C &= 3Nb \\ p_C &= \frac{a}{27b^2} \\ T_C &= \frac{8a}{27bk_B} \end{aligned} \quad (2.14)$$

Below the critical point, the liquid and the gas phases are separated by a first-order phase transition. The *equilibrium* transition between the two phases takes place when their chemical potentials and therefore the Gibbs free energies are equal.

Some comments are in place here.

- *Critical exponents.* The van der Waals equation was derived in a mean-field approximation, therefore all the exponents are mean-field like.
- *Liquid \rightarrow gas transition.* The liquid phase is stable until the point *A* is reached. *In equilibrium*, at this point, liquid starts to evaporate, the system enters the *coexistence region* (mixture of gas and liquid) until the point *B* is reached, which corresponds to the pure gas phase.

If equilibrium is not reached, the liquid phase persists beyond the point *A* towards the point *D*. In this region, the liquid is *metastable* and can (at least in principle) persist until the point *D* is reached. Beyond this point, the compressibility is negative, the liquid state is *unstable* (not accessible).

- *Gas \rightarrow liquid transition.* The situation is reversed. In equilibrium, the gas phase is stable until the point *B* is reached. This is the onset of the coexistence region which exists until the point *A* is reached. Metastable gas phase exists between the points *B* and *E* and, again, the region *E* to *D* is unstable.
- *Convexity.* As we have discussed in Section 2.3.3 already, the Helmholtz free energy *F* of a stable system has to be a concave function of *V*. For metastable states it is concave *locally* whereas for unstable states, it is convex.

- *Spinodals.* The transition from a metastable (homogeneous) state toward the equilibrium (mixed) state is connected with the *nucleation of droplets of the new phase*. Droplets first nucleate and then grow until the equilibrium (coexistence) is reached between the new and the old phase. Nucleation is a dynamic process.

The ultimate stability limit of the metastable state is called the spinodal line. Along this line the isothermal compressibility diverges, therefore the spinodal line is in a sense a line of critical points. In the van der Waals (mean-field) systems the separation between the metastable and unstable states is sharp. Beyond MFA, for systems with short-range correlations, however, the transition between metastable and unstable states is often gradual.

2.2 Ising Model

The Ising model is the prototype model for all magnetic phase transitions and is probably the most studied model of statistical physics. In this model, the spin operator \hat{S} is replaced by a *number*, which represents the z -component of the spin and is usually $S = \pm 1$ ("up" or "down"). The order parameter $m = \langle S \rangle$ is thus *one-dimensional*, it is a *two-component scalar*.

The Ising Hamiltonian is

$$\mathcal{H} = -\frac{1}{2} \sum_{i,j} J_{ij} S_i S_j - H \sum_i S_i \quad (2.15)$$

J_{ij} is the interaction energy between the spins i and j . Usually (not always) the sum is only over nearest neighbouring (NN) pairs, then $J_{ij} = J$. H is the external magnetic field (in energy units). The case $J > 0$ corresponds to ferromagnetism and $J < 0$ to antiferromagnetism. As we shall show later, the Ising model has a phase transition at finite T if $d > 1$. In $d = 1$ the transition is at $T = 0$. Therefore we say that $d = 1$ is the *lower critical dimension* of the Ising model.

In the low-temperature phase, the symmetry is broken, the spins are *ordered*, the order parameter is $m \neq 0$, whereas in the high-temperature phase the symmetry is not broken, m vanishes, the spins are disordered. Therefore this model is a prototype of all *order-disorder* transitions. Other examples of order-disorder transitions, I would like to mention, are lattice gases and binary alloys. All these systems are equivalent to each other, they all have the same critical behaviour (which - of course - depends on d), for given d they all belong to the same universality class.

2.2.1 Lattice Gas

Imagine that a fluid (gas or liquid) system is divided into a regular lattice of cells of volume roughly equal to the particle (atom, molecule) volume. We say that the cell is occupied if a particle centre falls into this cell. Since the cell volume is comparable to the particle volume, there can be not more than one particle per cell. Notice that microscopic details (like positions of the particles) are irrelevant in the critical region, therefore the lattice gas model is suitable for studying the critical behaviour of real gases although the space is discretized. (It is also assumed that the kinetic energy of the gas particles, which is neglected in the lattice gas model, does not contribute a singular part to the partition function.)

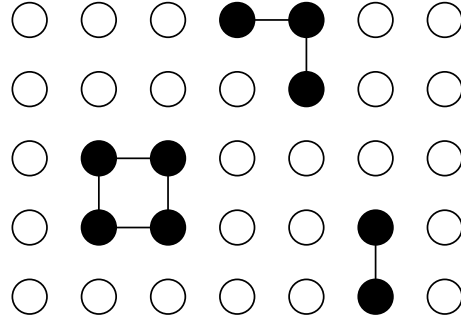


Figure 2.2: A configuration of a lattice gas in $d = 2$. Solid circles represent atoms and open circles empty sites. Lines show bonds between nearest-neighbouring atoms.

A lattice gas is thus a collection of particles whose kinetic energy is neglected and that are arranged into *discrete* cells. The cells are either occupied by one atom [closed circles in Fig. 2.2] or empty [open circles]. The Hamiltonian is:

$$\mathcal{H}_{lg} = \sum_{\langle i,j \rangle} u(r_{i,j}) p_i p_j - \mu \sum_i p_i, \quad (2.16)$$

where u is the interaction potential,

$$u(r_{i,j}) = \begin{cases} \infty & (r_{i,j} = 0) \\ -u & (r_{i,j} = \text{NN distance}) \\ 0 & (\text{otherwise}) \end{cases} \quad (2.17)$$

and $\langle i, j \rangle$ is the sum over pairs. p_i is the occupation number, it is 1 if the site is occupied by an atom, and zero otherwise. The occupation is, like the Ising spin,

a double-valued function. Therefore we expect that the two models are related. Indeed, if we set

$$p_i = \frac{1 + S_i}{2}, \quad (2.18)$$

the Hamiltonian 2.16 becomes:

$$\mathcal{H}_{lg} = -\frac{u}{4} \sum_{\langle i,j \rangle} S_i S_j - \left(\frac{\mu}{2} + \frac{uz}{4} \right) \sum_i S_i - \mathcal{H}_0. \quad (2.19)$$

Here, z is the number of NN neighbours (4 for square lattice, 6 for simple cubic lattice), and \mathcal{H}_0 is a constant, independent of S . Thus, we see that \mathcal{H}_{lg} is identical to the Ising Hamiltonian if we set $u/4 = J$ and $(\mu/2 + uz/4) = H$. For $u > 0$, the atoms attract and they "condense" at low temperatures whereas for $u < 0$ the atoms repel. At 1/2 coverage (one half of the lattice sites is occupied) and $u < 0$, the empty and occupied sites will alternate at low T , like spins in an antiferromagnet.

The lattice gas model in $d = 2$ is used to investigate, e.g., hydrogen adsorbed on metal surfaces. The substrate provides discrete lattice sites on which hydrogen can be adsorbed. By varying the pressure, the coverage (amount of adsorbed hydrogen) is varied. This corresponds to varying the field in magnetic systems.

There is an important difference between the lattice gas and the Ising models of magnetism. The number of atoms is constant (provided no atoms evaporate from or condense on the lattice), we say that the order parameter in the lattice gas models is *conserved* whereas it is not conserved in magnetic systems. This has important consequences for dynamics.

2.2.2 Binary Alloys

A binary alloy is composed of two metals, say A and B, arranged on a lattice. We introduce the occupations:

$$p_i = \begin{cases} 1 & \text{if } i \text{ is occupied by an A atom} \\ 0 & \text{if } i \text{ is occupied by a B atom} \end{cases}$$

$$q_i = \begin{cases} 1 & \text{if } i \text{ is occupied by a B atom} \\ 0 & \text{if } i \text{ is occupied by an A atom} \end{cases}$$

$$p_i + q_i = 1. \quad (2.20)$$

Depending on which atoms are nearest neighbours, there are three different NN interaction energies: u_{AA} , u_{BB} , and $u_{AB} = u_{BA}$. The Hamiltonian is:

$$\mathcal{H}_{ba} = -u_{AA} \sum_{\langle i,j \rangle} p_i p_j - u_{BB} \sum_{\langle i,j \rangle} q_i q_j - u_{AB} \sum_{\langle i,j \rangle} (p_i q_j + q_i p_j), \quad (2.21)$$

where each sum $\langle i, j \rangle$ runs over all NN pairs. If we write $p_i = (1 + S_i)/2$ and $q_i = (1 - S_i)/2$ then $S_i = 1$ if i is occupied by A and $S_i = -1$ for a B site. When there are 50% A atoms and 50% B atoms, $\sum_i S_i = 0$, and the Hamiltonian becomes an Ising Hamiltonian in the absence of a magnetic field:

$$\mathcal{H}_{ba} = -J \sum_{\langle i,j \rangle} S_i S_j + \mathcal{H}_0. \quad (2.22)$$

\mathcal{H}_0 is a constant, and $J = (u_{AA} + u_{BB} - 2u_{AB})/4$. For binary alloys, the order parameter is also conserved.

2.2.3 Ising model in $d = 1$

It is not difficult to solve the Ising model in the absence of H in $d = 1$ exactly. The Hamiltonian of a linear chain of N spins with NN interactions is

$$\mathcal{H}_{1d} = -J \sum_i S_i S_{i+1}. \quad (2.23)$$

We shall use *periodic boundary conditions*, that means that the spins will be

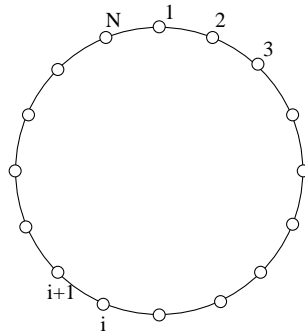


Figure 2.3: Lattice sites of an Ising model in one dimension with periodic boundary conditions.

arranged on a ring, see Fig. 2.3. Then, the N -th and the first spins are NN and the system is periodic. The partition function is:

$$Z = \sum_{\{S\}} e^{-\beta \mathcal{H}(\{S\})} = \sum_{\{S\}} e^{\beta J \sum_i S_i S_{i+1}} = \sum_{\{S\}} \prod_{i=1}^N e^{\beta J S_i S_{i+1}}. \quad (2.24)$$

(Don't forget: $\sum_{\{S\}}$ is the sum over all spin *configurations* – there are 2^N spin configurations – whereas \sum_i is the sum over all *sites*!)

The exponent in (2.24) can be written as:

$$e^{\beta J S_i S_{i+1}} = \cosh \beta J S_i S_{i+1} + \sinh \beta J S_i S_{i+1} = \cosh \beta J + S_i S_{i+1} \sinh \beta J. \quad (2.25)$$

The last equality holds because $S_i S_{i+1}$ can only be ± 1 and \cosh is an even function whereas \sinh is an odd function of the argument. The partition function is:

$$Z = (\cosh \beta J)^N \sum_{\{S\}} \prod_{i=1}^N [1 + K S_i S_{i+1}], \quad (2.26)$$

where $K = \tanh \beta J$. We work out the product and sort terms in powers of K :

$$\begin{aligned} Z &= (\cosh \beta J)^N \sum_{\{S\}} (1 + S_1 S_2 K)(1 + S_2 S_3 K) \cdots (1 + S_N S_1 K) \\ &= (\cosh \beta J)^N \sum_{\{S\}} [1 + K(S_1 S_2 + S_2 S_3 + \cdots + S_N S_1) + \\ &\quad + K^2(S_1 S_2 S_2 S_3 + \cdots) + \cdots + K^N(S_1 S_2 S_2 \cdots S_N S_N S_1)]. \end{aligned} \quad (2.27)$$

The terms, linear in K contain products of two different (neighbouring) spins, like $S_i S_{i+1}$. The sum over all spin configurations of this product vanish, $\sum_{\{S\}} S_i S_{i+1} = 0$, because there are two configurations with parallel spins ($S_i S_{i+1} = +1$) and two with antiparallel spins ($S_i S_{i+1} = -1$). Thus, the term linear in K vanishes after summation over all spin configurations. For the same reason also the sum over all spin configurations, which appear at the term proportional to K^2 , vanish. In order for a term to be different from zero, all the spins in the product must appear twice (then, $\sum_{\{S_i\}} S_i^2 = 2$). This condition is fulfilled only in the last term, which – after summation over all spin configurations – gives $2^N K^N$. Therefore the partition function of the Ising model of a linear chain of N spins is:

$$Z = (2 \cosh \beta J)^N [1 + K^N]. \quad (2.28)$$

In the thermodynamic limit, $N \rightarrow \infty$, $K^N \rightarrow 0$, and the partition function becomes:

$$Z(H = 0) = (2 \cosh \beta J)^N. \quad (2.29)$$

The free energy,

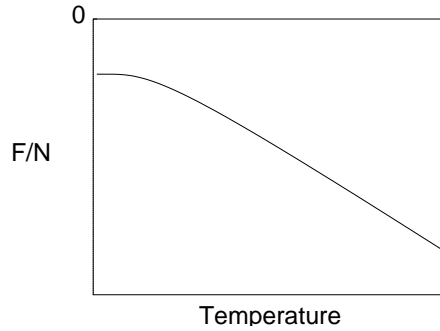


Figure 2.4: Free energy of the Ising model in $d = 1$.

$$F = -Nk_B T \ln(2 \cosh \beta J), \quad (2.30)$$

is regular at all finite T , therefore the Ising model in $d = 1$ has no phase transition at any finite T .

To investigate the possibility of a phase transition at $T = 0$, we will calculate the correlation function Γ and the correlation length ξ . We first replace J by J_j (at the end we put $J_j = J$), so that

$$Z = 2^N \prod_j \cosh \beta J_j. \quad (2.31)$$

The nearest-neighbour spin-spin correlation function is:

$$\begin{aligned} \Gamma(i, i+1) &= \langle S_i S_{i+1} \rangle = \frac{1}{Z} \sum_{\{S\}} S_i S_{i+1} e^{\beta \sum_j J_j S_j S_{j+1}} \\ &= \frac{1}{\beta Z} \frac{\partial Z}{\partial J_i} = \tanh(\beta J_i). \end{aligned} \quad (2.32)$$

The spin correlation function for arbitrary distance k ,

$$\Gamma(i, i+k) = \langle S_i S_{i+k} \rangle$$

$$\begin{aligned}
&= \langle S_i S_{i+1} S_{i+1} S_{i+2} S_{i+2} S_{i+k-1} S_{i+k-1} S_{i+k} \rangle \\
&= \frac{1}{Z} \frac{\partial}{\partial(\beta J_i)} \frac{\partial}{\partial(\beta J_{i+1})} \frac{\partial}{\partial(\beta J_{i+2})} \cdots \frac{\partial}{\partial(\beta J_{i+k-1})} Z, \quad (2.33)
\end{aligned}$$

$$\Gamma(k) \equiv \Gamma(i, i+k) = [\tanh(\beta J)]^k, \quad (2.34)$$

is just the product of all the nearest-neighbour correlation functions between the sites i and $i+k$, and is (if all J 's are equal) independent of the site i . The decay of the spin correlation function with distance is shown in Fig. 2.5. The correlations decay exponentially with distance and the correlation length is:

$$\frac{\xi}{a} = -\frac{1}{\ln \tanh \beta J}, \quad (2.35)$$

where a is the separation between the NN spins ("lattice constant"). In the limit $T \rightarrow 0$, the correlation length diverges exponentially as $T \rightarrow 0$:

$$\frac{\xi}{a} \approx \frac{1}{2} e^{2\beta J}. \quad (2.36)$$

Equations 2.34 and 2.36 tell us that there is no long range order at any finite T and

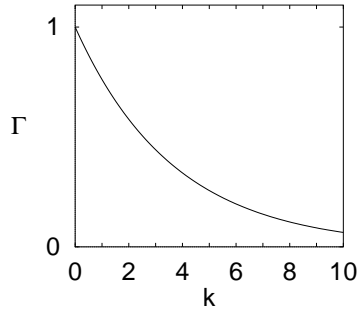


Figure 2.5: Decay of the correlations in the $d = 1$ Ising model.

that $T = 0$ is the critical temperature of the $d = 1$ Ising model. Such a behaviour is typical for systems in the lower critical dimension:

$$d > d_C : \quad T_C \neq 0 \quad (2.37)$$

$$d \leq d_C : \quad T_C = 0, \quad (2.38)$$

therefore we conclude that $d = 1$ is the *lower critical dimension* of the Ising model.

The zero-field isothermal susceptibility of the $d = 1$ Ising model is, according to the fluctuation-dissipation theorem,

$$\chi_T(H = 0) = \frac{\beta}{N} \sum_{i,j} \langle S_i S_j \rangle. \quad (2.39)$$

In the thermodynamic limit, χ_T becomes:

$$\begin{aligned} \chi_T(H = 0) &= \frac{\beta}{N} \sum_i \sum_{k=-N/2}^{N/2} \langle S_i S_{i+k} \rangle \\ &= \frac{\beta}{N} \sum_i \left[2 \sum_{k=0}^{N/2} \langle S_i S_{i+k} \rangle - \langle S_i S_i \rangle \right]. \end{aligned} \quad (2.40)$$

In the thermodynamic limit ($N \rightarrow \infty$), $\chi_T(H = 0)$ becomes

$$\chi_T(H = 0) = \beta \left[\frac{2}{1 - \tanh \beta J} - 1 \right] = \frac{1}{k_B T} e^{2J/k_B T}. \quad (2.41)$$

The susceptibility has an essential singularity of the type $\exp(1/T)$ as $T \rightarrow T_C = 0$ when $J > 0$. When $J < 0$, i.e., for antiferromagnetic NN interaction, the spin correlation function alternates in sign and the susceptibility is finite at all T .

The internal energy is:

$$E = -\partial \ln Z / \partial \beta = -JN \tanh \beta J. \quad (2.42)$$

Notice the similarity between the internal energy and the short-range correlation function $\Gamma(i, i + 1)$! This is a general property of the internal energy.

The entropy is $S = (E - F)/T$,

$$S(H = 0) = Nk_B \ln [2 \cosh \beta J] - \frac{NJ}{T} \tanh \beta J. \quad (2.43)$$

In the limit $T \rightarrow 0$, $S \rightarrow 0$ and in the limit $T \rightarrow \infty$, $S \rightarrow Nk_B \ln 2$, see Fig. 2.6.

The heat capacity $C_H = T(\partial S / \partial T)_H$ is:

$$C_H = Nk_B \left(\frac{\beta J}{\cosh \beta J} \right)^2. \quad (2.44)$$

The heat capacity is finite at all temperatures (no singularity at $T = 0$!), see Fig. 2.7.

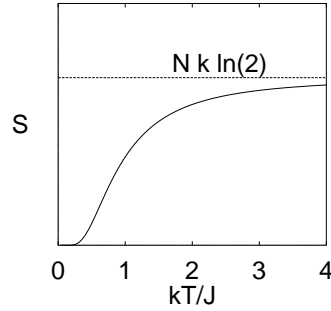


Figure 2.6: Temperature dependence of the entropy for the zero-field $d = 1$ Ising model.

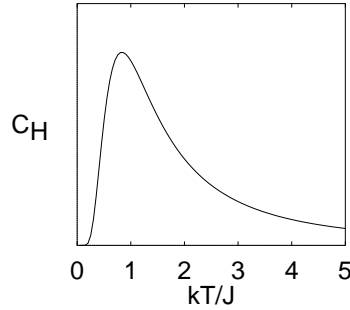


Figure 2.7: Temperature dependence of the heat capacity for the zero-field $d = 1$ Ising model.

2.2.4 Ising Model in $d = 2$

The Ising model for $H = 0$ in $d = 2$ dimensions has been solved exactly by Onsager. The calculation is lengthy, therefore we will first show that the Ising model in $d = 2$ has a phase transition at finite T and then quote the Onsager's result.

Consider a two-dimensional square lattice with all the spins on the outer boundary "up". This simulates an arbitrarily weak external field and breaks the up-down symmetry of the model. To show that spontaneous ordering exists, it is enough to show that in the thermodynamic limit at a sufficiently small but finite temperature there are more spins "up" than "down". Now we will estimate the fraction of "down" spins from the average size of a spin-down domain.

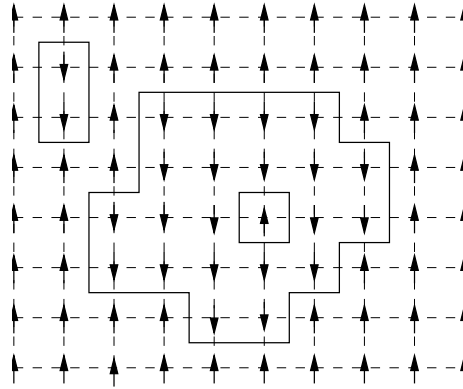


Figure 2.8: A possible domain configuration of the Ising model on a square lattice.

With the above boundary conditions, all the domain walls must be closed loops, see Fig. 2.8. In a wall the neighbouring spins are antiparallel, this costs an energy $2J$ per wall segment and the probability, that a domain with its total wall length (perimeter) b is thermally excited, is given by the Boltzmann factor $\exp[-2\beta Jb]$. For large b , there are many possible domain wall configurations with the same wall length b . For an estimate of the number m of possible wall configurations, assume that on each site the wall can either be straight or make a kink to either side. This makes a factor of 3 for each wall segment (except for the last one where it is bound to close the loop). Besides, a wall can start on each lattice site, this makes a factor of N . Therefore, $m(b)$ is limited by:

$$m(b) \leq 3^{b-1} N \quad (2.45)$$

and the average number of different thermally excited domains with perimeter b is:

$$\leq 3^{b-1} N e^{-2\beta Jb}. \quad (2.46)$$

The number of "down" spins in a domain with perimeter b is $N_- \leq (b/4)^2$ and the average concentration of all "down" spins is:

$$\begin{aligned} \frac{\langle N_- \rangle}{N} &\leq \sum_{b=4,6,8,\dots} \frac{b^2}{16} 3^{b-1} e^{-2\beta Jb} \\ &= \frac{1}{48} \frac{\partial^2}{\partial \alpha^2} \sum_{b=4,6,8,\dots} 3^b e^{-\alpha b} \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{48} \frac{\partial^2}{\partial \alpha^2} \sum_{x=2,3,4,\dots} (9e^{-2\alpha})^x \\
&= \frac{1}{48} \frac{\partial^2}{\partial \alpha^2} \left(\frac{q^2}{1-q} \right) \\
&= \frac{q^2(4-3q+q^2)}{12(1-q)^3}, \tag{2.47}
\end{aligned}$$

where $\alpha = 2\beta J$, $x = 2b$, $q = 9\exp(-2\alpha)$. For large, finite β , q is small and the above ratio is $< 1/2$. This means that the order parameter $\langle N_+ - N_- \rangle / N$ is greater than zero at low T and the Ising model in $d = 2$ (and of course also in $d > 2$) has a phase transition at finite T .

The exact critical temperature of the (anisotropic) Ising model on a rectangular lattice is given by the equation:

$$\sinh\left(\frac{2J_x}{k_B T_C}\right) \sinh\left(\frac{2J_y}{k_B T_C}\right) = 1,$$

where J_x and J_y are the spin interaction energies in x and y directions, respectively.

2.3 Potts and Related Models

In the previous Section we have discussed different classical two-state models (which are all mapped on the Ising model). A large class of classical physical systems cannot be described by the Ising model. As an example consider adsorption of noble gases on surfaces, like krypton on graphite. Graphite provides a two-dimensional hexagonal lattice of adsorption sites. Kr is large and once it occupies one hexagon, it is unfavourable for another Kr atom to occupy any of its NN hexagons, see Fig. 2.9. At full coverage, thus, Kr forms a regular lattice with $1/3$ of hexagons covered by Kr. Since there are three entirely equivalent positions (states!) for Kr (denoted by A, B, and C), the system is described by a three-state model. The adsorbed Kr forms A,B,C *domains*, separated by *domain walls*. It is clear from Fig. 2.9 that the A-B and the B-A walls are not equivalent, they have different energies in general.

In the Potts model there are q states available to each site. The Potts Hamiltonian (with only NN interactions) is:

$$\mathcal{H} = -J \sum_{\langle i,j \rangle} \delta_{\sigma_i, \sigma_j}. \tag{2.48}$$

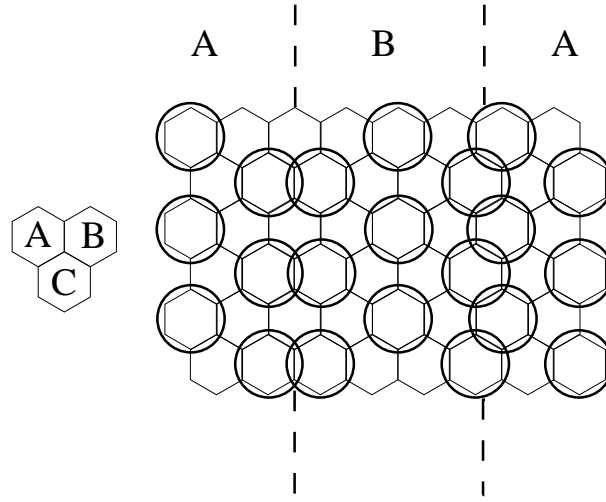


Figure 2.9: Kr on graphite. There are three equivalent adsorption sites, but Kr atoms are too big to occupy all of them – only one third of the sites (i.e., either "A", "B", or "C" sites) can be occupied. We talk about "A," "B," or "C" domains. The domains are separated by domain walls. The A-B domain wall has a different structure (and energy) than the B-A wall.

J is the interaction energy and $\delta_{\sigma_i, \sigma_j}$ is the δ -function which is = 1 if the NN sites i and j are in the same *state* and zero otherwise,

$$\delta_{\sigma_i, \sigma_j} = \begin{cases} 1 & \sigma_i = \sigma_j \text{ with } (i, j) \in NN \\ 0 & \text{otherwise} \end{cases} \quad (2.49)$$

$\sigma_i = 1, 2, \dots, q$ is a number which describes the *state* of the site i . Thus, in the Potts model, the interaction energy is equal to $-J$ if the two NN sites are in the same state and vanishes otherwise. For $q = 2$, the Potts model is equivalent to the Ising model ($S = \pm 1$) because each site has two possible states in both cases. The two states can be visualized as two points on a line (which is a one-dimensional object). The mapping from the Potts to the Ising model is accomplished by writing:

$$\delta_{S_i S_j} = \frac{1 + S_i S_j}{2}. \quad (2.50)$$

For $q = 3$ the three states of the Potts model can be visualized as corners of a triangle (which is a two-dimensional object) whereas for $q = 4$, the four states

belong to corners of a tetrahedron, which is a 3-dimensional object. We conclude that the order parameter of the q -state Potts model is a q -component scalar. In $d = 2$, the Potts model has a continuous phase transition for $q \leq 4$ and a first-order transition for $q > 4$.

Order parameter. Let x_A, x_B, \dots, x_q be the fractions of cells (sites) that are in the states A, B, \dots , q , respectively. Of course, $x_A + x_B + \dots + x_q = 1$. In the disordered phase, all the fractions are equal whereas in the ordered phase one state, say A, will have greater probability than the other, $x_A > (x_B, x_C, \dots, x_q)$. We make a simplified approach, assuming that $x_B = x_C = \dots = x_q$. The order parameter m measures the asymmetry in occupancies,

$$m = x_A - \frac{x_B + \dots + x_q}{q-1}$$

$$x_A = \frac{1 + (q-1)m}{q}, \quad x_B = x_C = \dots = x_q = \frac{1-m}{q}$$

It is easy to verify that for $m_A = 0$: $x_A = x_B = x_C = \dots = x_q$ and for $m = 1$: $x_A = 1$ and $x_B = x_C = \dots = x_q = 0$. thus, m_A is the order parameter which is $\neq 0$ in the ordered phase and vanishes in the disordered phase.

I would like to mention another generalization of the 2-state model, that is the q -state (chiral) clock model. In this model, the states are imagined as hands of a clock (i.e., vectors on a circle) that can take only discrete orientations, $\Theta_i = (2\pi/q)\sigma_i$ (where $\sigma_i = 0, 1, 2, \dots, q-1$). The q -state chiral clock Hamiltonian is

$$\mathcal{H} = -J \sum_{\langle i,j \rangle} \cos \left(\Theta_i - \Theta_j + \vec{\Delta} \cdot \vec{r}_{i,j} \right) \quad (2.51)$$

Δ is the chirality which breaks the symmetry of the model, it produces an asymmetry in the interaction between the NN sites. The interaction energies between, say, $\sigma_i = 0, \sigma_j = 1$ and $\sigma_i = 1, \sigma_j = 0$ are different. The two-state clock model is identical to the Ising model and, if $\Delta = 0$, the 3-state clock model is identical to the 3-state Potts model. For $q > 3$, the clock and the Potts model are not identical. The two-dimensional clock and Potts models are widely used in surface physics. The above mentioned Kr on graphite can be described with the 3-state chiral clock model with $\Delta \neq 0$.

2.4 Two-dimensional xy Model

A very interesting situation appears when the spin is a continuous vector in a plane (dimensionality of the order parameter is 2) and the spatial dimensionality is also

$d = 2$. As we have seen, the Ising model in $d \geq 2$ does have a phase transition at finite T_C . In the xy model, however, spin excitations (spin waves) are easier to excite thermally than for the Ising model, they are so strong that they even destroy long-range order at any finite T (Mermin-Wagner theorem). On the other hand, there has been evidence that some kind of a transition does take place at a finite T_C . What is the ordering and what is the nature of this transition? This question was answered in the early seventies by Berezinskii and by Kosterlitz and Thouless.

The Hamiltonian of the xy model is:

$$\mathcal{H} = -J \sum_{\langle i,j \rangle} \vec{S}_i \cdot \vec{S}_j = -J \sum_{\langle i,j \rangle} \cos(\Phi_i - \Phi_j), \quad (2.52)$$

where the spins are classical vectors of unit length aligned in the plane. Φ_i is the angle between the spin direction and an arbitrarily chosen axis and is a continuous variable. For simplicity, let us consider the square lattice and let the sum be only over nearest neighbours. Only slowly varying configurations ($\Phi_i - \Phi_j \ll 1$) give considerable contributions to the partition function, therefore we expand \mathcal{H} :

$$\begin{aligned} \mathcal{H} - E_0 &= \frac{J}{2} \sum_{\langle i,j \rangle} (\Phi_i - \Phi_j)^2 = \frac{Ja^2}{2} \sum_i |\vec{\nabla} \Phi(i)|^2 \\ &= \frac{J}{2} \int d^2r |\vec{\nabla} \Phi(r)|^2. \end{aligned} \quad (2.53)$$

Here a is the lattice constant and we transformed the last sum over all lattice sites into an integral over the area. $\Phi(r)$ is now a scalar *field*.

Possible excitations above the (perfectly ordered) ground state are spin waves. They are responsible for destroying long-range order at any finite T . We will not consider spin waves here. Another possible excitation above the ground state is a vortex, shown in Fig. 2.10. Any line integral along a closed path around the centre of the vortex will give

$$\oint \vec{\nabla} \Phi(\vec{r}) \cdot d\vec{r} = 2\pi q \quad \Rightarrow \quad \begin{cases} \Phi = q\phi \\ |\vec{\nabla} \Phi| = q/r \end{cases} \quad (2.54)$$

where q is the vorticity (q is integer; $q = 1$ for the vortex shown in Fig. 2.10) and ϕ is the polar angle to \vec{r} . From the equations (2.53) and (2.54) we find the energy of this vortex:

$$E = \int_a^R 2\pi r dr \frac{J}{2} \left(\frac{q}{r} \right)^2 = \pi J \ln \left(\frac{R}{a} \right) q^2. \quad (2.55)$$

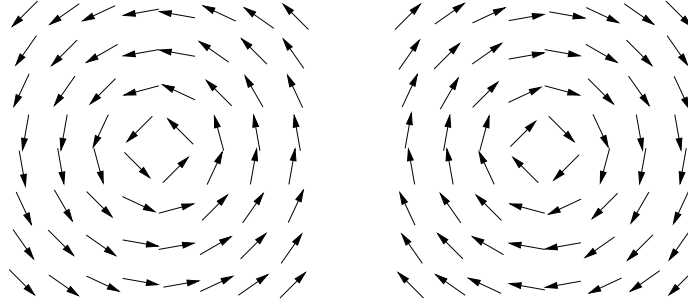


Figure 2.10: Two examples of a vortex with vorticity $q = 1$.

R is the size of the system. Notice that the energy of the vortex increases logarithmically with the size of the system. In estimating the entropy associated with a vortex we realize that a vortex can be put on each lattice site therefore the number of configurations is about $(R/a)^2$ and the entropy of one vortex is

$$S \approx 2k_B \ln \left(\frac{R}{a} \right). \quad (2.56)$$

The free energy, $F = E - TS$, is positive and large at low temperatures. That means that it is not very easy to (thermally) excite single vortices. As temperature increases, the entropy term causes F to decrease. F vanishes (changes sign) at a critical temperature:

$$k_B T_C = \frac{\pi J}{2}. \quad (2.57)$$

Above T_C a large number of vortices is thermally excited.

However, this does not mean that there are no vortices at low T . The easiest way to see this is to draw an analogy to a system of electrical charges in two dimensions. If \vec{k} is a unit vector perpendicular to the plane of the system, then $\vec{\mathcal{E}} = \vec{\nabla}(q\phi) \times \vec{k}$ is equal to the electric field produced by a charge of strength $2\pi\epsilon_0 q$ positioned at the centre of the vortex. The energy of an electric field is

$$E_{el} = \frac{\epsilon_0}{2} \int d^2 r |\vec{\mathcal{E}}|^2 = \frac{\epsilon_0}{2} \int d^2 r |\vec{\nabla}(q\phi) \times \vec{k}|^2, \quad (2.58)$$

whereas the energy of a system of vortices is

$$E_{vort} = \frac{J}{2} \int d^2 r |\vec{\nabla} \Phi \times \vec{k}|^2 = \frac{J}{\epsilon_0} E_{el}. \quad (2.59)$$

The energy of a system of vortices of strength q_i is equivalent to the energy of a system of electric charges! There is almost a complete analogy between a system of vortices and the two-dimensional Coulomb gas. We have seen that, like the electrostatic energy of a single charge in two dimensions, also the energy of a vortex diverges logarithmically with the size of the system. However, a pair of opposite charges, i.e., a dipole does have a finite electrostatic energy because at large distances the fields of two opposite charges almost cancel. In analogy to the energy of two charges $\pm q$ at \vec{r}_1 and \vec{r}_2 we write the total energy of a system of two opposite vortices:

$$E_{2 \text{ vort}} = 2\pi J q^2 \ln \left(\frac{|\vec{r}_1 - \vec{r}_2|}{a} \right). \quad (2.60)$$

This self-energy of a vortex pair (of a dipole) is very small when the vortices are close together. Thus, it is difficult to excite a single vortex but it is very easy to excite a pair of opposite vortices. Whereas a single vortex influences the spin configuration on the whole lattice, the effect of a pair of opposite vortices cancels out on large distances from the pair so that the effect is limited to an area of the order $|\vec{r}_1 - \vec{r}_2|^2$.

Since the size of the system does not appear in (2.60), but it does appear in the expression for the entropy, we will have a finite density of bound vortex pairs. The equilibrium density of vortex pairs will depend on the pair-pair (dipole-dipole in the Coulomb gas case) interaction which we neglected here.

Now we have the following picture. At low but > 0 temperature ($T < T_C$) there will be a finite density of low-energy vortex *pairs* ("dipoles") of zero total vorticity. At high temperatures ($T > T_C$), on the other hand, there will be a large concentration of *single* vortices. The critical temperature, we estimated above, is thus associated with *unbinding (dissociation) of vortex pairs*.

Very interesting is the behaviour of the spin correlation function $\langle \vec{S}_i \cdot \vec{S}_j \rangle$. We already mentioned that spin-wave excitations break the long-range order. If we neglect the spin wave excitations (which have nothing to do with the phase transition, they do not behave in a critical way at any finite temperature), then:

$$\lim_{|\vec{r}_i - \vec{r}_j| \rightarrow \infty} \langle \vec{S}_i \cdot \vec{S}_j \rangle = \begin{cases} > 0 & \text{for } T < T_C \\ 0 & \text{for } T > T_C. \end{cases} \quad (2.61)$$

This follows from the fact that bound vortices don't influence the spin orientations at large distances whereas they do if they are free.

In a more realistic approach, one must take into account interaction between the vortex pairs, which was neglected in (2.60). Surrounding vortices (dipoles!) will screen the (electric) field and will thus reduce the self-energy of a vortex.

A sophisticated real-space renormalization group treatment of the two-dimensional xy model [Kosterlitz 1974] shows that the correlation length ξ diverges above T_C as:

$$\xi \propto \exp(b/t^{1/2}) \quad b \approx 1.5 \quad (2.62)$$

where $t = (T - T_C)/T_C$. Below T_C , however, the correlation length is always infinite and the correlations decay algebraically (like in a critical point)! Therefore we can say that we have a line of critical points extending from zero to T_C . The phase transition is characterized by a change from algebraic (power-law) decay of correlations below T_C to an exponential decay above T_C .

Analogous to vortex unbinding transition in the xy model is the unbinding of opposite charges of a dipole in the two-dimensional Coulomb gas. At low temperatures, the Coulomb gas is composed of neutral dipoles, bound pairs, and is not conducting. At high temperature it is composed of free $+$ and $-$ charges which are conducting.

Another analogy to the xy model are the dislocation lines in a (three-dimensional) crystal. At low temperatures, pairs of dislocation lines with opposite Burgers vectors can exist. At T_C , the dislocations dissociate. This is the basis of the dislocation theory of melting. In this latter example, stiffness can play the role of the order parameter, it is $\neq 0$ below the melting temperature and vanishes in the liquid phase.

Let me stress again that the described behaviour of Coulomb gases or xy systems is specific only for *two-dimensions*!

2.5 Spin Glasses

The (vitreous) glasses are amorphous materials, their lattice is not crystalline but is strongly disordered. Topological defects destroy the long-range order in the crystalline structure whereas the short-range order is still preserved to a high degree. Example: topological defects on a square lattice would be triangles or pentagons. Of course, the defects cause also a deformation of the squares.

Spin glasses are the magnetic (spin) analogue of the (vitreous) glasses. In magnetic systems, the exchange interaction is often an oscillating function of distance between the spins. If such a system is diluted with non-magnetic ions, the magnetic atoms are randomly positioned on the lattice and thus experience random exchange interactions with other, neighbouring atoms. Example: $\text{Au}_{1-x}\text{Fe}_x$, where x is small.

Another system, which behaves in a similar way, is a random mixture of a ferromagnetic and an antiferromagnetic materials. The exchange interaction between nearest neighbours i and j is random, it depends on the type of the magnetic atoms i and j .

A Hamiltonian which reflects the essential features of a spin glass is:

$$\mathcal{H} = - \sum_{i,j} J(R_{i,j}) \vec{S}_i \cdot \vec{S}_j - \sum_i \vec{H} \cdot \vec{S}_i. \quad (2.63)$$

$J(R_{i,j})$ is a random variable and the first sum includes also more distant pairs than just nearest neighbours. Random exchange interactions cause that – even at low temperatures – the spins are not ordered, the system has *no long-range order in the ground state*. The “usual” long-range order parameter

$$m = \frac{1}{N} \sum_i \langle S_i \rangle \quad (2.64)$$

vanishes for a spin glass. Yet, in the experiment, a broad maximum in the specific heat and a rather sharp maximum in the zero-field susceptibility, connected with hysteresis and remanescence, are seen. So, there must be a kind of a transition as the temperature is varied, we shall call this a spin–glass transition and the low-temperature state the spin-glass phase.

At high temperatures, the spins behave like in a normal paramagnet, they flip around dynamically so that the thermodynamic average $\langle S_i \rangle$ vanishes. There is no long-range order in the system and also the local spontaneous magnetization, i.e., the (time) average of the spin at a site i vanishes, $\langle S_i \rangle = 0$ for each site. At low temperatures, however, the spins *freeze* in a disordered configuration, the system is in a state where $\langle S_i \rangle \neq 0$ although the average magnetization still vanishes, $m = 0$. The spin-glass transition is thus a freezing transition.

Something happens to the dynamics of the spins, therefore we must consider time-dependent spins, $S(t)$. To distinguish between the two phases, we introduce the (Edwards-Anderson) order parameter:

$$q = \lim_{t \rightarrow \infty} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_i \langle S_i(t_0) S_i(t_0 + t) \rangle. \quad (2.65)$$

q clearly vanishes in the paramagnetic phase and is > 0 in the spin-glass phase, see Fig. 2.12. It measures the mean square local spontaneous magnetization. Notice that freezing is a gradual process, which starts at T_f where the first spins freeze and ends at $T = 0$ where all the spins are frozen.

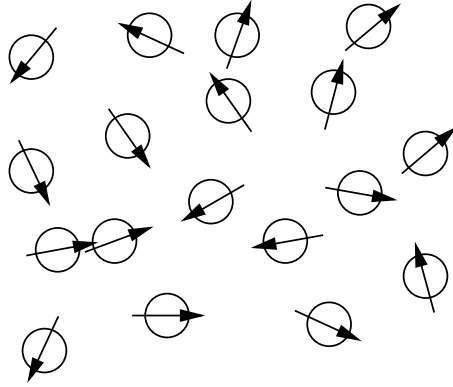


Figure 2.11: Spins in the low-temperature, spin-glass phase are frozen in a disordered configuration.

Since, in the spin-glass phase the spins are (at least partially) frozen in a disordered configuration, we can easily imagine that some spins are frustrated and that the spins do not freeze in a unique way. Freezing is history dependent, the configuration into which the spins freeze, depends on the conditions (e.g, on the magnitude of the magnetic field) close to the freezing temperature. If we, upon cooling, apply an external field pointing "up", there will be more spins frozen in the "up" orientation and will stay in this orientation even if we switch the field off in the low-temperature phase. This means that the systems can fall into a metastable configuration from which it is not able to escape if the temperature is low. The free energy, thus, has many, many local minima in the *configuration space*, see Fig. 2.13. In the thermodynamic limit ($N \rightarrow \infty$), some hills between the local minima will become infinitely high and the system will be trapped in such a minimum forever. In the time average, thus, the system will explore only a part of the total configuration space. This means that the *time average* of a physical quantity is no longer equal to the *ensemble average* (where one explores the whole configuration space). A system in the *spin-glass phase is thus not ergodic* in the thermodynamic limit! Since it is ergodic in the paramagnetic phase, we say that at the freezing transition the ergodicity breaks down.

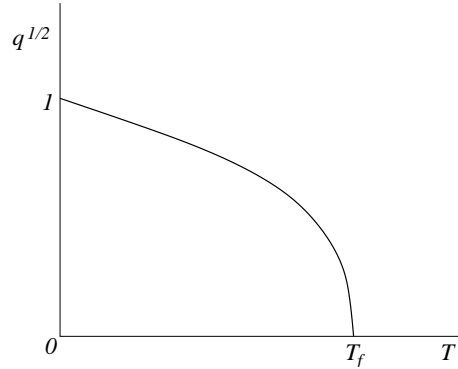


Figure 2.12: Temperature dependence of the order parameter q . Upon cooling, the spins start to freeze at T_f ; below T_f , more and more spins are frozen until, eventually, at $T = 0$ all the spins are frozen.

2.5.1 Some Physical Properties of Spin Glasses

The static susceptibility, Fig. 2.14, below the freezing temperature depends strongly on the way the experiment is performed. $\chi(T)$ is largest and roughly temperature independent in a "field cooling" (FC) cycle, i.e., if the field H is applied above T_f and then the sample is cooled at constant H below T_f . This susceptibility is reversible, it is independent of the history. In contrast, the "zero-field cooled" (ZFC) susceptibility is obtained if the sample is cooled to $T \approx 0$ at $H = 0$ and then the sample heated in the field. The difference between χ_{FC} and χ_{ZFC} confirms the conjecture that the system can be frozen in different spin configurations (which, of course, also have different susceptibilities).

The conjecture that freezing takes place at T_f is confirmed also by the *ac* susceptibility measurements, shown in Fig 2.15 for $\text{Cu}_{1-x}\text{Mn}_x$ ($x = 0.9\%$) at different frequencies. The lower the frequency, the more pronounced (less rounded) the cusp is. We see also that the position of the cusp shifts to lower temperatures as the frequency decreases. So, the freezing temperature should be defined as the maximum of $\chi(\nu)$ in the limit as $\nu \rightarrow 0$.

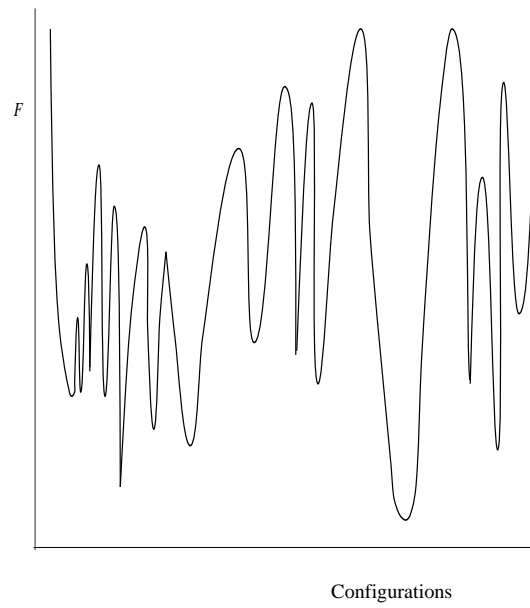


Figure 2.13: Free energy landscape in the configuration space of a spin glass at low temperatures. There is one global and many local minima.

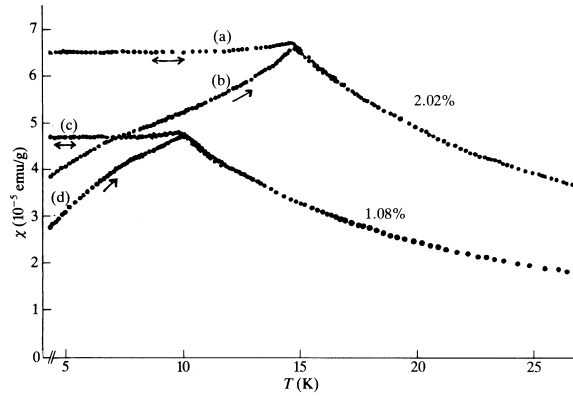


Figure 2.14: Temperature dependence of the static susceptibility of $\text{Cu}_{1-x}\text{Mn}_x$ for $x = 1.08$ and 2.02% . (b) and (d): cooled to low temperature at $H = 0$ and then heated at $H = 5.9$ Oe; (a) and (c): cooled and/or heated in $H = 5.9$ Oe. [From Nagata et al., Phys. Rev. **B 19**, 1633 (1979).]

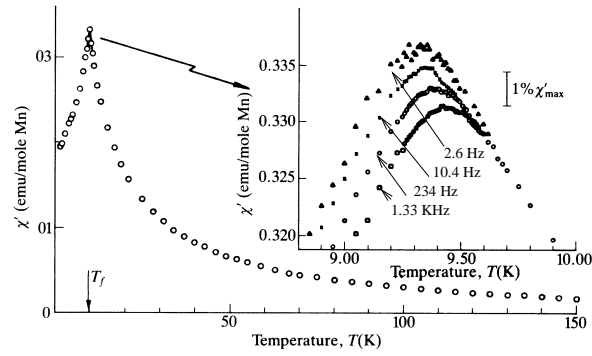


Figure 2.15: Temperature dependence of the frequency-dependent susceptibility of $\text{Cu}_{1-x}\text{Mn}_x$ ($x = 0.9\%$). [From C.A.M. Mulder et al., Phys. Rev. **B** **23**, 1384 (1982), *ibid.* **25**, 515 (1982).]

Chapter 3

Quantum Models

In classical statistical mechanics we did not know the exact microscopic state of the system (location in the phase space). We made use of the ergodic hypothesis and replaced the time average of a physical quantity with an ensemble average, i.e., an average over many equivalent systems.

In quantum statistical mechanics the situation is similar.

3.1 Quantum Statistics

In this Chapter we will first introduce the density matrix and then make an introduction to the Bose condensation.

In *quantum mechanics*, the energy levels are discrete and the state of the system is described by a wavefunction. Let us first assume that we know the dynamical state (wavefunction) $|\psi\rangle$ of the whole system and the *complete* set of stationary states $|n\rangle$ which are the eigenstates of the full Hamiltonian of the compound system,

$$\mathcal{H}|n\rangle = E_n|n\rangle. \quad (3.1)$$

(We assume that the states $|n\rangle$ are normalized.) n denotes the quantum states (e.g., momenta) of all particles in the system. E_n is the total energy of all particles of the system in the compound state n . Any dynamical state of the system can be expressed as a linear combination of these eigenstates,

$$|\psi(t)\rangle = \sum_n c_n(t)|n\rangle, \quad (3.2)$$

The coefficients $c_n(t)$ define a point in the (Hilbert) space of wavefunctions $|n\rangle$. The (time-dependent) wavefunction $|\psi(t)\rangle$ which obeys the Schrödinger equation

$$i\hbar \frac{d|\psi\rangle}{dt} = \mathcal{H}|\psi\rangle \quad (3.3)$$

is thus the quantum-mechanical analogue of the point in phase space of classical statistical mechanics. The expectation value of a physical quantity A (which does not necessarily commute with \mathcal{H}) is:

$$\langle A(t) \rangle = \langle \psi(t) | A | \psi(t) \rangle = \sum_{m,n} c_n^* c_m \langle n | A | m \rangle \quad (3.4)$$

The last expression holds for any complete set of states $|n\rangle$, also if they are not the eigenstates of the full Hamiltonian.

3.1.1 The Density Matrix

As the number of particles in the system increases, the separation between the energy levels decreases rapidly, and, in the thermodynamic limit, it becomes extremely high. On the other hand, the energy levels are never completely sharp. They are broadened for many reasons, including the uncertainty principle. In a macroscopic system (with extremely dense energy levels), the levels will thus always overlap. Consequently, a macroscopic system will never be in a strictly stationary quantum state, it will always be in a *dynamical* state, in a time-dependent *mixture* of stationary states.

It is neither feasible nor possible to find a complete description of such a system. Like in *classical statistical mechanics*, we first assume that the system is *ergodic*, we replace the time average by an ensemble average over many quantum systems at the same instance. We talk about probabilities of finding the system in the states $|n\rangle$. The dynamical state is no longer a unique quantum state but is a *statistical mixture of quantum states*. We introduce the *density matrix* as the ensemble average of $c_m^* c_n$:

$$\begin{aligned} \rho &= |m\rangle \rho_{m,n} \langle n| \\ \rho_{m,n} &\equiv \langle m | \rho | n \rangle = \frac{1}{\Omega} \sum_{i=1}^{\Omega} c_n^i c_m^i. \end{aligned} \quad (3.5)$$

The sum over i runs over Ω configurations of the system, $|\psi^i\rangle$, all being part of the same ensemble, and $c_m^i = \langle m | \psi^i \rangle$. We do an ensemble average because our information on the system is not complete.

The density matrix is a statistical operator with the following important properties:

- The density matrix completely defines an ensemble of quantum systems and it carries all the information that is available for a quantum statistical ensemble.
- Its diagonal elements $\rho_{n,n}$ tell the probability that the system is in the state $|n\rangle$ whereas its non-diagonal elements $\rho_{m,n}$ tell the probability of a transition from the state $|n\rangle$ to the state $|m\rangle$. For stationary states, thus, ρ has to be diagonal, it commutes with the Hamiltonian.
- The density matrix is normalized,

$$\text{Tr}(\rho) = 1. \quad (3.6)$$

The trace runs over any complete set of states $|n\rangle$ and is independent of the choice of the basis set.

- The (ensemble and quantum-mechanical) average value $\langle A \rangle$ of any physical quantity A is equal to:

$$\langle A \rangle = \langle \psi | A | \psi \rangle = \sum_{n,m} \rho_{m,n} \langle n | A | m \rangle = \text{Tr}(\rho A) = \text{Tr}(A \rho). \quad (3.7)$$

The trace in quantum statistical mechanics plays the role of integration over the whole available phase space in classical statistical mechanics.

- The dynamics of the system is completely described by:

$$i\hbar \frac{\partial \rho}{\partial t} = [\mathcal{H}, \rho]. \quad (3.8)$$

This is the “Schrödinger equation” of the density matrix.

3.1.2 Ensembles in Quantum Statistical Mechanics

Now, we shall specify the density matrix and its relation to thermodynamic quantities.

- *Quantum Microcanonical ensemble.*

The energy of systems belonging to a microcanonical ensemble is fixed, therefore it is convenient to use the energy eigenstates as a basis. Then the density matrix is diagonal. According to the postulate of equal weights, all the states $|n\rangle$ with the energy E_n between E and $E + \Delta E$ are equally probable and let there be $\Delta\Omega$ such states. For a quantum system, $\Delta\Omega$ is an integer number. The probability for the system being in one of these states is $p = 1/\Delta\Omega$ and the entropy is equal to:

$$S = -k_B \langle \ln p \rangle. \quad (3.9)$$

In general, we must use the density matrix instead of simple probabilities p and the entropy of a quantum system is:

$$S = -k_B \langle \ln \rho \rangle = -k_B \text{Tr}(\rho \ln \rho). \quad (3.10)$$

Again, the average $\langle \rangle$ means ensemble *and* quantum average – both averages cannot be separated.

- *Canonical ensemble.* In the energy representation ρ is diagonal and we can use the same arguments as in the case of classical canonical distributions. The matrix elements are equal to:

$$\rho_{m,m} = \frac{1}{Z} e^{-\beta E_m}, \quad (3.11)$$

where the normalization constant is (again) the partition function

$$Z = \sum_m e^{-\beta E_m}. \quad (3.12)$$

In an arbitrary basis, the density matrix can be written more generally as

$$\rho = \frac{1}{Z} e^{-\beta \mathcal{H}} \quad \text{with} \quad Z = \text{Tr} e^{-\beta \mathcal{H}} \quad (3.13)$$

(ρ and \mathcal{H} are operators!). Thus, for a canonical ensemble, the average (thermal and quantum- mechanical) of an operator A is:

$$\langle A \rangle = \text{Tr}(\rho A) = \frac{1}{Z} \text{Tr}(e^{-\beta \mathcal{H}} A) \quad (3.14)$$

In particular, the internal energy is

$$E = \langle \mathcal{H} \rangle = \frac{1}{Z} \text{Tr} (\mathcal{H} e^{-\beta \mathcal{H}}) = -\frac{\partial}{\partial \beta} \ln \text{Tr} e^{-\beta \mathcal{H}}$$

$$E = -\frac{\partial}{\partial \beta} \ln Z. \quad (3.15)$$

The free energy is

$$F = -k_B T \ln Z. \quad (3.16)$$

All the thermodynamic relations are the same as before (Tables I - III are still valid), the only difference is that now we have to calculate the partition function as the trace of the density matrix.

The entropy in a canonical ensemble, S , is the ensemble average over all (energy) states of the microcanonical entropy S :

$$S = -\langle k_B \ln \rho \rangle = -k_B \text{Tr}(\rho \ln \rho). \quad (3.17)$$

- *Grand canonical ensemble.* In the grand canonical ensemble the operator ρ operates on a generalized Hilbert space which is the direct sum of all Hilbert spaces with fixed number of particles. The density matrix is:

$$\rho = \frac{1}{\Xi} e^{-\beta(\mathcal{H} - \mu N)},$$

where the grand partition function is

$$\Xi(T, V, \mu) = \text{Tr} e^{-\beta(\mathcal{H} - \mu N)}.$$

Notice that now N is an operator in the generalized Hilbert space. As in the classical case, Ξ is equal to the canonical partition function Z , weighted by the fugacity $z = \exp(\beta\mu)$ and summed over N . The ensemble average of an operator A in the grand canonical ensemble is

$$\langle A \rangle = \frac{1}{\Xi} \text{Tr}(A e^{-\beta(\mathcal{H} - \mu N)}).$$

3.2 Bose Systems

3.2.1 Ideal Bose Gas

Now we shall consider the properties of an ideal gas of particles with integer spin and mass $m > 0$. As usually for ideal gases, we neglect the interaction between the particles. The Hamiltonian of the ideal Bose gas is:

$$\mathcal{H} = \sum_{\vec{p}} \frac{p^2}{2m} n_{\vec{p}}.$$

$n_{\vec{p}}$ is the number of particles in the state with momentum \vec{p} . Because we are dealing with Bosons, $n_{\vec{p}}$ can also be > 1 . We want to study condensation where two phases come into contact, the number of particles in one phase is not fixed, therefore we must work with the grand-canonical distribution. The grand-canonical partition function of ideal Bose gas is:

$$\Xi = \sum_{\{n_{\vec{p}}\}} e^{-\beta \sum_{\vec{p}} (p^2/2m) n_{\vec{p}} + \beta \mu N}, \quad N = \sum_{\vec{p}} n_{\vec{p}}.$$

(Here we will disregard the $(2S + 1)$ factor which comes from the spin degeneracy.) In the grand partition function, each sum over $n_{\vec{p}}$ is over all non-negative integers and, since the particles do not interact, we can split the partition function into a product:

$$\begin{aligned} \Xi &= \sum_{n_0=0}^{\infty} \sum_{n_1=0}^{\infty} \dots \left\{ [e^{\beta \mu}]^{n_0} [e^{-\beta(p_1^2/2m - \mu)}]^{n_1} [e^{-\beta(p_2^2/2m - \mu)}]^{n_2} \dots \right\} = \quad (3.18) \\ &= \prod_{\vec{p}} \left\{ \sum_n [e^{-\beta(p^2/2m - \mu)}]^{n_1} \right\} = \prod_{\vec{p}} \frac{1}{1 - z e^{-\beta p^2/2m}}. \end{aligned}$$

Here we have introduced the *fugacity* $z = e^{\beta \mu}$. The equation of state of an ideal Bose gas is (see Table I):

$$PV = k_B T \ln \Xi(T, V, \mu)$$

$$\frac{PV}{k_B T} = - \sum_{\vec{p}} \ln [1 - e^{-\beta(p^2/2m - \mu)}]$$

(3.19)

The thermal average of N is determined by the chemical potential μ (see Tables):

$$\langle N \rangle = -\frac{1}{\beta} \frac{\partial \ln \Xi}{\partial \mu} = \sum_{\vec{p}} \frac{1}{e^{\beta(p^2/2m - \mu)} - 1}. \quad (3.20)$$

We can write $\langle N \rangle = \sum_{\vec{p}} n_{\vec{p}}$ where

$$n_{\vec{p}} = \frac{1}{e^{\beta(p^2/2m - \mu)} - 1} \quad (3.21)$$

is the Bose-Einstein distribution function, which tells the occupation probability of a given (non-degenerate) state \vec{p} . For $\langle N \rangle$ to be finite and positive, the second term in the denominator has to be < 1 for any \vec{p} , in particular for $p = 0$. This means that μ must be < 0 and $0 < z < 1$.

For large V we replace the sum by an integral:

$$\sum_{\vec{p}} \rightarrow \frac{4\pi V}{h^3} \int_0^\infty p^2 dp \quad (3.22)$$

and we get:

$$\begin{aligned} \langle N \rangle &= \frac{4\pi V}{h^3} \int_0^\infty p^2 dp \frac{1}{e^{\beta(p^2/2m - \mu)} - 1} \\ &= V g_{3/2}(z) \left(\frac{mk_B T}{2\pi\hbar^2} \right)^{3/2} = \frac{V}{\lambda^3} g_{3/2}(z), \end{aligned} \quad (3.23)$$

where

$$g_\nu(z) = \frac{1}{\Gamma(\nu)} \int_0^\infty dx \frac{x^{\nu-1}}{e^x z^{-1} - 1} = \sum_{l=1}^\infty \frac{z^l}{l^\nu} \quad (3.24)$$

and λ is the thermal wavelength,

$$\lambda = \left(\frac{2\pi\hbar^2}{mk_B T} \right)^{1/2}.$$

The functions $g_{3/2}(z)$ and $g_{5/2}(z)$ are shown in Fig 3.1. For later use: $g_{3/2}(1) \approx 2.612$ and $g_{5/2}(1) \approx 1.342$. The pressure is:

$$P = \left(\frac{m}{2\pi\hbar^2} \right) (k_B T)^{5/3} g_{5/2}(z) = \frac{k_B T}{\lambda^3} g_{5/2}(z). \quad (3.25)$$

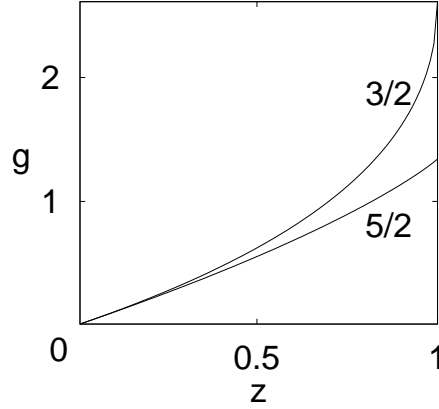


Figure 3.1: The functions $g_{3/2}(z)$ and $g_{5/2}(z)$.

Now let us start increasing the pressure at fixed (high enough) T . This increases the chemical potential, the fugacity, and the density of particles, $\langle N \rangle / V$. However, we know that z must be < 1 . What happens when $z \rightarrow 1$? Can we further increase the pressure or the density of particles?

To answer this question we must go back to Eq. (3.20). The term with $\vec{p} = 0$ which was not included in (3.23) (The density of states $n(E)$ has a \sqrt{E} dependence on E and vanishes as $E \rightarrow 0$.) becomes singular (divergent) when $\mu \rightarrow 0$ and it has to be treated separately. This divergence has very important consequences for a Bose gas, as we shall show now. Instead of Eq. (3.20) we must write:

$$\langle N \rangle = \sum_{\vec{p} \neq 0} \frac{1}{e^{\beta(p^2/2m - \mu)} - 1} + \frac{z}{1 - z} = \frac{V}{\lambda^3} g_{3/2}(z) + \frac{z}{1 - z} \quad (3.26)$$

and the equation of state is:

$$\frac{PV}{k_B T} = -\frac{4\pi V}{h^3} \int_0^\infty p^2 dp \ln [1 - z e^{-\beta p^2/2m}] - \ln(1 - z)$$

$$\frac{P}{k_B T} = \frac{1}{\lambda^3} g_{5/2}(z) - \frac{1}{V} \ln(1 - z),$$

(3.27)

The last terms in Eqs. (3.26) and (3.27) come from the term $p = 0$ and correspond to the particles in the lowest energy, in the ground state. In the thermodynamic limit the last term in (3.27) vanishes whereas it remains finite in (3.26). We will denote

$$\langle N \rangle = \langle N' \rangle + \langle N_0 \rangle, \quad (3.28)$$

where $\langle N' \rangle$ is the average number of bosons in the continuum (excited, $p > 0$) states,

$$\langle N' \rangle = \frac{V}{\lambda^3} g_{3/2}(z) \quad (3.29)$$

and $\langle N_0 \rangle$ the average number of bosons in the ground state,

$$\langle N_0 \rangle = \frac{z}{1-z}. \quad (3.30)$$

The average density of bosons in the continuum states reaches its maximum value when $g_{3/2}(z)$ is maximal, that is for $z \rightarrow 1$ ($\mu \rightarrow 0$),

$$\frac{\langle N'_{max}(T) \rangle}{V} = g_{3/2}(1) \left(\frac{mk_B T}{2\pi\hbar^2} \right)^{3/2} \propto T^{3/2}. \quad (3.31)$$

The fact that $\langle N' \rangle$ is limited and cannot exceed $\langle N'_{max} \rangle$ has its origin in the quantum-mechanical and statistical nature of the system. The number of available states in a box with volume V is finite and is equal to $\frac{V}{\lambda^3}$. Each state, on the average, accommodates $n_{\vec{p}}$ particles, where $n_{\vec{p}}$ is given by the Bose-Einstein distribution function, Eq. (3.21). If there are more than $\langle N' \rangle$ particles in the system, they are *pushed* into the ground state.

3.2.2 Bose Einstein Condensation

We will now investigate the behaviour of *ideal* Bose gases, described by Eqs. (3.26, 3.27). Let us consider a system of N bosons in a volume V .

At high enough temperature, $\langle N'_{max} \rangle$ is so large that $N < \langle N'_{max} \rangle$ and the chemical potential is determined from

$$N = \langle N' \rangle = V \left(\frac{mk_B T}{2\pi\hbar^2} \right)^{3/2} g_{3/2}(z). \quad (3.32)$$

This means that $z < 1$ and $\mu < 0$. The density of particles in the ground state is

$$\langle N_0 \rangle / V = z / (1 - z) / V \rightarrow 0 \quad \text{as} \quad V \rightarrow \infty. \quad (3.33)$$

Almost all the particles are in the $|\vec{p}| > 0$ states and the ground state is macroscopically empty. The equation of state simplifies to

$$\frac{P}{k_B T} = \frac{g_{5/2}(z)}{\lambda^3} \quad (3.34)$$

At very high T , $g_{3/2}(z)$ and $g_{5/2}(z)$ are $\ll 1$ and we approximate [see the series for g_n , Eq. (3.24)]:

$$g_{5/2}(z) \approx g_{3/2} \approx z \quad (3.35)$$

and the equation of state becomes:

$$\frac{P}{k_B T} = g_{5/2}(z) \left(\frac{mk_B T}{2\pi\hbar^2} \right)^{3/2} \approx z \left(\frac{mk_B T}{2\pi\hbar^2} \right)^{3/2} = \frac{z}{\lambda^3}. \quad (3.36)$$

z is determined by the density of particles:

$$\frac{N}{V} = g_{3/2}(z) \left(\frac{mk_B T}{2\pi\hbar^2} \right)^{3/2} \approx z \left(\frac{mk_B T}{2\pi\hbar^2} \right)^{3/2} = \frac{z}{\lambda^3}. \quad (3.37)$$

The fugacity $z = \langle N \rangle \lambda^3 / V$ vanishes as $T^{-3/2}$ at high T and for fixed density. This justifies the approximation (3.35). Eliminating z from the last two equations yields the equation of state of the ideal Bose gas at high T :

$PV = \langle N \rangle k_B T$

(3.38)

which is identical to the equation of state of a classical ideal gas (as it should be at high T !).

Now we lower the temperature by keeping N and V fixed. Eq. (3.32) tells us that $g_{3/2}(z)$ must increase. This means that z and μ also increase. Eventually, a temperature is reached, where z reaches its maximum value ($z \rightarrow 1$, $N = \langle N'_{max} \rangle$). This is the transition temperature T_C and is determined by the condition

$$\frac{N}{V} = g_{3/2}(1) \left(\frac{mk_B T_C}{2\pi\hbar^2} \right)^{3/2} \quad (3.39)$$

from which we find:

$$T_C = \frac{2\pi\hbar^2}{k_B m [g_{3/2}(1)]^{2/3}} \left(\frac{N}{V}\right)^{2/3}. \quad (3.40)$$

If we insert this expression into the equation of state, Eq. (3.34), we obtain the pressure at which the transition takes place:

$$P_C = k_B T_C^{5/2} g_{5/2}(1) \left(\frac{mk_B}{2\pi\hbar^2}\right)^{3/2} \propto T_C^{5/2}. \quad (3.41)$$

At T_C , the ensemble average of the density of particles in the continuum states reaches its maximum value whereas the ground state is still empty, $\langle N_0 \rangle / V = 0$.

Below T_C , $\langle N'_{max} \rangle$ further decreases and becomes $< N$ (we keep N fixed). The chemical potential further approaches zero (μ cannot be exactly $= 0$, because that would mean that $N = \infty$) and $z \rightarrow 1$ in such a way that the total number of particles is kept constant. z is very close to 1, we can put $g_n(z) \approx g_n(1)$ and the number of particles in the continuum states is

$$\langle N' \rangle = V \left(\frac{mk_B T}{2\pi\hbar^2}\right)^{3/2} g_{3/2}(1). \quad (3.42)$$

When $\langle N' \rangle < N$, the ground state starts to fill, $\langle N_0 \rangle$ increases. Below T_C , thus, only a part of the particles can be accommodated in the continuum states. The rest must go into the ground state! The number of bosons in the ground state is:

$$\langle N_0 \rangle = N - \langle N' \rangle = N - V \left(\frac{mk_B T}{2\pi\hbar^2}\right)^{3/2} g_{3/2}(1)$$

from which we get:

$$\frac{\langle N_0 \rangle}{N} = 1 - \left(\frac{T}{T_C}\right)^{3/2}. \quad (3.43)$$

$\langle N_0 \rangle / N$ is finite, the state with $\vec{p} = 0$ is occupied with a *macroscopic* number of particles, the particles *condense* in the *momentum* space into the zero-momentum

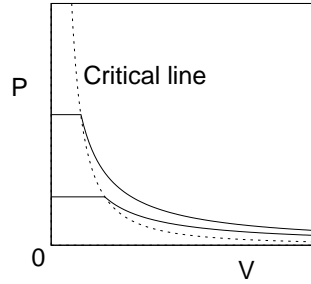
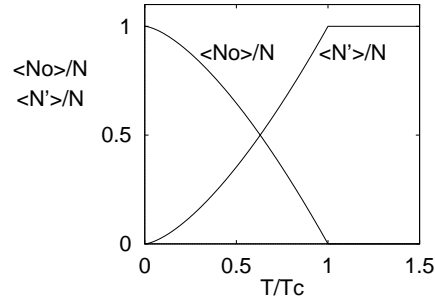


Figure 3.2: Isotherms of the ideal Bose gas.

Figure 3.3: Temperature dependence of the relative number of bosons in the ground ($\langle N_0 \rangle/N$) and excited ($\langle N' \rangle/N$) states.

state. This is the *Bose-Einstein condensation*. The fugacity is, from Eq. (3.30): $z = \langle N_0 \rangle / (\langle N_0 \rangle + 1)$. In the thermodynamic limit, when V and $\langle N_0 \rangle \rightarrow \infty$, $z = 1$ and $\mu = 0$ below T_C . The equation of state is:

$$P = g_{5/2}(1) \left(\frac{mk_B T}{2\pi\hbar^2} \right)^{3/2} k_B T = \frac{g_{5/2}(1)}{g_{3/2}(1)} \frac{\langle N' \rangle}{V} k_B T. \quad (3.44)$$

Only the particles in the continuum states (in the gas phase) contribute to pressure. The particles in the ground state (condensate) are at rest, they cannot exert any pressure. If the density is increased, the extra particles fall into the ground state and the pressure does not increase. At fixed $T < T_C$ the density of particles in the gas phase is constant and P is independent of V , see Fig. 3.2. The temperature dependence of $\langle N_0 \rangle$ is shown in Fig. 3.3. In the limit $T \rightarrow 0$, $\langle N'_{max} \rangle \rightarrow 0$ and all the particles are in the ground state, $N = \langle N_0 \rangle$.

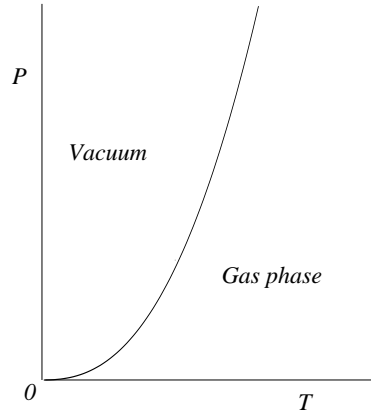


Figure 3.4: Phase diagram of an ideal Bose gas.

The corresponding *phase diagram* is shown in Fig. 3.4. The normal, gas phase exists for $T > T_C$, i.e., to the right of the transition line. On the line, given by (3.41), condensation takes place and since the particles in the ground state don't contribute to pressure, the condensate lies on the transition line itself. Notice that the "critical points" in Bose systems form a *line* in the P vs. T or in the P vs. V planes and are not in a single point like in van der Waals gases.

We now invert Eq. (3.44),

$$V = \frac{g_{5/2}(1)}{g_{3/2}(1)} \frac{\langle N' \rangle}{P} k_B T. \quad (3.45)$$

As $T \rightarrow 0$, also $\langle N' \rangle \rightarrow 0$ and therefore $V \rightarrow 0$ at any finite pressure. The condensed phase does not occupy any volume. This means that the ideal Bose gas can be compressed to zero volume without any increase in pressure.

Some comments: The condensate (ground state) contributes neither to E , nor to C_V , P , or V ! At low T , C_V vanishes like $T^{3/2}$ in contrast to the photon gas, where $C_V \propto T^3$. They differ because the density of states is different in both cases. There are more excited states available for particles than for photons and the specific heat is greater.

Of course, this, and infinite compressibility are the artifacts of the ideal gas model. In reality, because of atomic repulsion, the volume of the condensed phase does not vanish and the compressibility does not diverge. The equation of state of a non-ideal Bose gas is shown in Fig. 3.5.

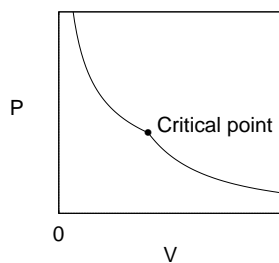


Figure 3.5: An isotherm of a non-ideal Bose gas.

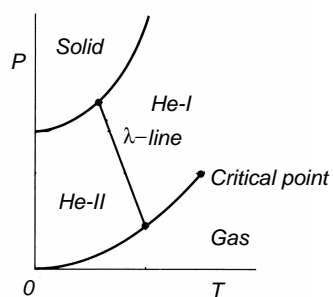


Figure 3.6: Phase diagram of ^4He . He I is the normal, and He II the superfluid phase.

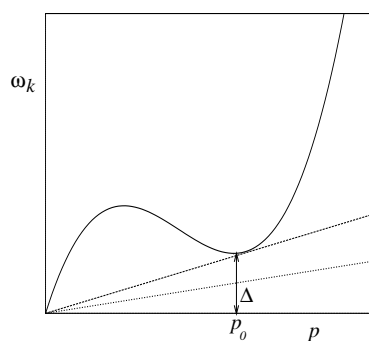


Figure 3.7: Dispersion curve of elementary excitations in liquid ^4He . The excitations around the minimum at p_0 are called rotons.

3.3 Liquid ^4He

A candidate for the Bose-Einstein condensation at low T is liquid ^4He . The phase diagram is shown in Fig. 3.6. He I is the normal liquid and He II exhibits superfluid behaviour. From the fact that the line separating the solid and the fluid phases is horizontal at $T \rightarrow 0$, we conclude with the help of the Clausius-Clapeyron equation,

$$\frac{dp}{dT} = 0 = \frac{S_{\text{fluid}} - S_{\text{solid}}}{V_{\text{fluid}} - V_{\text{solid}}} \quad (3.46)$$

that liquid He II at $T \rightarrow 0$ has no more entropy than solid helium, which is highly ordered, of course. We conclude that the particles in the ground state ($T \rightarrow 0$) have already the highest possible degree of order although they are in the fluid and not in the solid phase. We interpret the He II phase as a mixture of two fluids, a normal fluid with finite viscosity and a super-fluid (formed of particles in the ground state) with zero viscosity. Characteristic for the superfluid state is also a very large heat conductivity (as a consequence, He II does not boil, the atoms only evaporate from the surface). What is the ground state of He II and what are the excitations?

We shall first discuss possible excitations above the ground state of liquid ^4He and only then the properties of the ground state. In the liquid, the atoms are close together, they are strongly interacting (exchanging momenta). A state of an atom with fixed momentum is not an eigenstate of the Hamiltonian, the atoms are subject to a *collective motion*. The elementary excitations in the system are collective excitations, *quasiparticles*, *phonons* (longitudinal because it is a liquid) and *rotons* (see Fig. 3.7). At low momenta, the excitations are phonons with a linear dispersion relation,

$$\epsilon_p = cp. \quad (3.47)$$

At high momenta (short wavelengths), the interaction between the atoms causes a minimum in the dispersion curve. Elementary excitations around this minimum are (for historical reasons) called rotons. Their energy is written in the form:

$$\epsilon_p = \Delta + \frac{(p - p_0)^2}{2m^*}, \quad (3.48)$$

where Δ is the energy gap, $\Delta/k_B = 8.6\text{K}$, and $m^* = 1.0 \times 10^{-27}\text{kg}$ is the effective mass of rotons. $p_0 = 2.0 \times 10^{-24}\text{kg m/s}$. At low T , the density of these quasiparticles is low and they don't interact.

The grand canonical partition function for phonons ($\mu = 0$) is, in analogy with (3.18):

$$\begin{aligned}\ln \Xi_{ph} &= -\frac{4\pi V}{h^3} \int_0^\infty p^2 dp \ln(1 - e^{-\beta cp}) \\ &= \frac{4\pi V}{3(h\beta c)^3} \int_0^\infty \frac{x^3 dx}{e^x - 1} \\ &= \frac{4\pi^5 V}{45(hc)^3} (k_B T)^3\end{aligned}\tag{3.49}$$

The grand canonical partition function for rotons (also $\mu = 0$) is:

$$\ln \Xi_{rot} = -\frac{4\pi V}{h^3} \int_0^\infty p^2 dp \ln(1 - e^{-\beta\Delta - \beta(p-p_0)^2/2m^*})$$

At low T (large β), $e^{-\beta\Delta - \beta(p-p_0)^2/2m^*}$ is small and we approximate $\ln(1 - x) \approx -x$. (This is equivalent to using the Maxwell-Boltzmann distribution.)

$$\begin{aligned}\ln \Xi_{rot} &= \frac{4\pi V}{h^3} e^{-\beta\Delta} \int_0^\infty p^2 dp e^{-\beta(p-p_0)^2/2m^*} \\ &\approx \frac{4\pi V}{h^3} e^{-\beta\Delta} \int_{-\infty}^\infty p^2 dp e^{-\beta(p-p_0)^2/2m^*} \\ &= \frac{4\pi V}{h^3} \sqrt{\frac{2\pi m^*}{\beta}} \left(p_0^2 + \frac{m^*}{\beta} \right) e^{-\beta\Delta}\end{aligned}\tag{3.50}$$

From these equations, one can calculate the specific heat (and other thermodynamic quantities) at low temperatures.

3.3.1 Superfluidity

Consider an isolated system (in which energy and momentum are conserved) composed of a long cylindrical capillary filled with liquid ^4He at $T = 0$. Initially, let the capillary tube rotate along the symmetry axis with the tangential velocity v_0 and let the fluid be at rest. We would expect that - because of friction - energy would be transferred from the tube to the fluid and that the fluid will gradually be excited out of the ground state. Let us check under which conditions the energy and momentum can be transferred to the fluid. Immediately after the tube started to rotate, the tube had the kinetic energy $Mv_0^2/2$ (M is the mass of the tube) and the fluid was at rest and had the ground-state energy E_0 . At a later time, we expect the tube velocity to decrease by Δv (because we consider an isolated system) and

that in the fluid an (at least one) elementary excitation with energy ϵ_p and momentum \vec{p} is created. The conservation laws demand that the energy at a later time is:

$$\frac{1}{2}M(v_0 - \Delta v)^2 + \epsilon_p + E_0 = \frac{1}{2}Mv_0^2 + E_0 \quad (3.51)$$

and that the momentum is:

$$M(v_0 - \Delta v) + p_t = Mv_0 \quad \Rightarrow \quad M\Delta v = p_t \quad (3.52)$$

where p_t is the excitation momentum in the tangential direction. For $\Delta v \neq 0$, the first equation gives the inequality

$$\epsilon_p = Mv_0\Delta v - \frac{1}{2}M(\Delta v)^2 < Mv_0\Delta v. \quad (3.53)$$

After eliminating $M\Delta v$, the conservation laws tell us that

$$\vec{v} \cdot \vec{p} > \epsilon_p. \quad (3.54)$$

Inspection of Fig. 3.7 shows that this is only possible if v is larger than a threshold velocity v_c . At small relative velocity between the tube and the fluid, the rotons are not excited, the momentum is not transferred to the fluid, the fluid will stay at rest. This means that there is no friction, no viscosity - this is the superfluid phase. (As we will see later, roton excitations are not responsible for the excitation out of the ground state. Vortices have higher energy, but lower ϵ/p ratio.)

3.3.2 Vortex Excitations

There are other collective excitations in liquid ^4He that have higher energies than phonons or rotons, but which nonetheless have much lower energy to momentum ratio than rotons. The creation of these excitations allows energy and momentum to be transferred to the fluid at a much lower critical velocity than in the case of rotons.

Let us now discuss such excitations of the condensate. The wave function of an excited collective state is

$$\Psi = \psi e^{i\phi(r)}, \quad (3.55)$$

the (mass) current density is

$$\vec{j} = \frac{i\hbar}{2} [\Psi \vec{\nabla} \Psi^* - \Psi^* \vec{\nabla} \Psi] = \hbar \psi^2 \vec{\nabla} \phi(\vec{r}). \quad (3.56)$$

On the other hand,

$$\vec{j} = \frac{N}{V} m \vec{v} \quad (3.57)$$

where $N/V = \rho = \psi^2$. From the two equations we get for the particle velocity

$$\vec{v} = \frac{\hbar}{m} \vec{\nabla} \phi(\vec{r}). \quad (3.58)$$

The wavefunction must be single-valued and the phase ϕ is determined modulo 2π , therefore the closed-path integral of the momentum is

$$m \oint \vec{v} \cdot d\vec{l} = \hbar \oint \vec{\nabla} \phi d\vec{l} = nh, \quad (3.59)$$

where h is the Planck's constant and n an arbitrary integer.

To understand the nature of these excited states it is best to consider liquid ^4He within a small rotating bucket. If the bucket is slowly brought into rotation, the liquid will behave in the following way. At small angular velocities of the bucket, the liquid will remain at rest. If the angular velocity is increased beyond some threshold angular velocity, the energy is transferred to the fluid, and the fluid will make an abrupt transition to a state in which there is a *vortex line* running from top to bottom, parallel to the axis of rotation. The vortex line looks like a small whirlpool, the fluid flows around the line with the velocity that decreases with distance from the line. The *vorticity*, defined as the line integral of the momentum along any path encircling the vortex line, Eq. (3.59), is quantized. As the angular velocity is increased, more and more quantized vortex lines are created. Once a vortex line is created, the macroscopically occupied single-particle ground state is no longer a zero-momentum eigenstate. It is a quantum state with a flow pattern that includes the vortex lines. When there are many vortex lines, they will distribute themselves in a pattern so that the average fluid velocity (averaged over many vortices) is the same as in the case of rigid-body rotation.

To summarize, the quantization of vortices is related to the phase factor of the condensate wave function, defined in Eq. (3.60). In the lowest-energy state, ψ and ϕ do not depend on \vec{r} , they are constant. The state in which ϕ varies with \vec{r} is an excited state which describes the fluid with one or several vortices.

3.3.3 Order Parameter

A superfluid can be considered as a mixture of two coexisting fluids, of a normal component and of a superfluid component. At a first glance, one would choose

the density of particles in the ground (superfluid) state as the order parameter, because it is $\neq 0$ in the low-temperature phase and vanishes at T_C . However, in an interacting system, the operator N_0 doesn't commute with \mathcal{H} , the single-particle momenta are not good quantum numbers. On the other hand we already know that the ground state is a state with zero momentum $\vec{p} = 0 = -i\hbar\vec{\nabla} \Rightarrow \vec{\nabla} = 0$. This means that the *wavefunction* is a constant in the ground state. Therefore we choose for the order parameter the *statistical average of the ground-state wave function* which commutes with \mathcal{H} and which can be written in the form:

$$\langle \Psi \rangle = \frac{1}{\Xi} \text{Tr} [\Psi e^{-\beta(\mathcal{H}-\mu N)}] = \psi e^{i\phi} \quad (3.60)$$

(ψ is the amplitude and ϕ the phase of the wavefunction). The density of the superfluid component is $\langle N_0 \rangle / V = \psi^2 = \text{constant}$. In the superfluid phase, a *macroscopically occupied single-particle quantum-mechanical ground state with zero momentum extends through the whole volume available to the fluid. Each atom is not at any particular place, it is simultaneously everywhere in the macroscopic volume!* This is the meaning of the order parameter defined in (3.60).

Symmetry breaking. There is a strong parallel between $\langle \Psi \rangle$ and the spontaneous magnetization $\langle \vec{M} \rangle$ in ferromagnets,

$$\langle \vec{M} \rangle = \frac{1}{Z} \text{Tr} (\vec{M} e^{-\beta \mathcal{H}}), \quad (3.61)$$

where \vec{M} is the (operator for the) total magnetic moment and

$$\mathcal{H} = -J \sum_{\langle i,j \rangle} \vec{S}_i \cdot \vec{S}_j.$$

In the absence of an external magnetic field, thus, \mathcal{H} is invariant under rotations, \vec{M} and $-\vec{M}$ occur with equal probability and therefore $\langle \vec{M} \rangle$ is always zero. But we know that the ferromagnets do have spontaneous magnetization at low T . The resolution lies in "spontaneous symmetry breaking" - the ground state of a Hamiltonian does *not* possess the full symmetry of the Hamiltonian. The ground state wavefunction is not rotationally invariant, because $\langle \vec{M} \rangle$ points along a definite direction in space whereas \mathcal{H} is rotationally invariant. In the case of ferromagnets, thus, the transition corresponds to spontaneous breaking of the rotational symmetry. One must therefore redefine $\langle \vec{M} \rangle$ in such a way that $+\vec{M}$ and $-\vec{M}$ are not both included in the partition function with equal weights. This is most easily done by imposing an infinitesimal external field H :

$$\langle \vec{M} \rangle = \lim_{H \rightarrow 0} \lim_{V \rightarrow \infty} \frac{1}{Z} \text{Tr} M e^{-\beta(\mathcal{H} - \vec{M} \cdot \vec{H})}. \quad (3.62)$$

Returning to Bose systems, the Hamiltonian has a *global gauge invariance*, the wave functions are invariant with respect to the phase shift,

$$\Psi(\vec{r}) \rightarrow e^{i\phi(\vec{r})} \Psi(\vec{r}) \quad (3.63)$$

where ϕ is an arbitrary phase factor (real number). Bose condensation corresponds to a *spontaneous breaking of global gauge invariance*. When calculating the order parameter, we imagine, like in magnetic systems that the system is subjected to an (infinitesimal) external field η which couples linearly to the order parameter. We calculate the statistical average of $\Psi(\vec{r})$ in the thermodynamic limit:

$$\langle \Psi(\vec{r}) \rangle = \lim_{\eta \rightarrow 0} \lim_{V \rightarrow \infty} \frac{1}{\Xi} \text{Tr} \Psi(\vec{r}) e^{-\beta \mathcal{E}}, \quad (3.64)$$

where

$$\mathcal{E} = \mathcal{H} - \mu N - \int d^3r [\Psi(\vec{r}) \eta(\vec{r}) + \Psi^\dagger(\vec{r}) \eta^\dagger(\vec{r})]. \quad (3.65)$$

The only essential difference with the ferromagnetic case is that the external field here is fictitious, it cannot be realized experimentally.

Chapter 4

Methods of Statistical Mechanics

4.1 Mean-Field Theories

Very few models of statistical mechanics have been solved exactly; in most of the cases one has to rely on approximative methods. Among them, the mean-field approximation (MFA) is one of the most widely used. The advantage of the mean-field theory is its simplicity and that it correctly predicts the qualitative features of a system in most cases.

The essence of the mean-field theory is the assumption of statistical independence of the local ordering (spins in the case of magnetic systems). The interaction terms in the Hamiltonian are replaced by an effective, "mean field" term. In this way, all the information on correlations in the fluctuations is lost. Therefore, the mean-field theory is usually inadequate in the critical regime. It usually gives wrong critical exponents, in particular at low spatial dimensions when the number of NN is small.

The MFA becomes exact in the limit as the number of interacting neighbours $z \rightarrow \infty$. This is the case when the range of interaction $z \rightarrow \infty$ or if the spatial dimension is high. In both cases, a site "feels" contributions from many neighbours and therefore the fluctuations average out or become even irrelevant (for Ising systems, they become irrelevant in $d \geq 4$). At low d , the MFA must be used with caution. Not only that it predicts wrong critical exponents, its predictions are even qualitatively wrong sometimes. For example, it predicts long-range order and finite critical temperature for the $d = 1$ Ising model. In the following, we shall introduce and formulate the MFA for magnetic systems. In the Appendix A we summarize basic thermodynamic relations for magnetic systems. The results,

however, are valid also for other systems.

Two formulations of the mean-field approximation. We will now formulate the mean-field approximation in two different ways.

- A: The “Weiss molecular field theory” according to its “inventor” Pierre Weiss. This method is straightforward and easy to understand. It gives us an expression for the order parameter but does not tell us anything about the free energy or partition function.
- B: The second method will be the Bragg-Williams approximation which is more sophisticated and is based on the free energy minimization.

We will develop the mean-field approximation on the Ising model in an external magnetic field.

4.1.1 Weiss Molecular Field of an Ising System

Weiss formulated a theory of ferromagnetism in which he assumed that the effect of the neighbouring spins on a given spin can be described by a fictitious “molecular field” which is proportional to the average magnetization of the system.

We start with the Ising Hamiltonian in an external magnetic field:

$$\mathcal{H} = -J \sum_{\langle ij \rangle} S_i S_j - H \sum_i S_i. \quad (4.1)$$

The local field acting on the spin at the site i is:

$$H_i = J \sum_j S_j + H. \quad (4.2)$$

It depends on the configuration of all the neighbouring spins around the site i . If the lattice is such that each spin has many neighbours, then it is not a bad approximation to replace the actual value of the neighbouring spins by the *mean value* of all spins,

$$\sum_j S_j \rightarrow zm, \quad (4.3)$$

where z is the number of nearest neighbours and m is the mean (average) magnetization per site, $m = \frac{1}{N} \sum_j S_j$. With (4.3), the local field becomes

$$H_i = zJm + H, \quad (4.4)$$

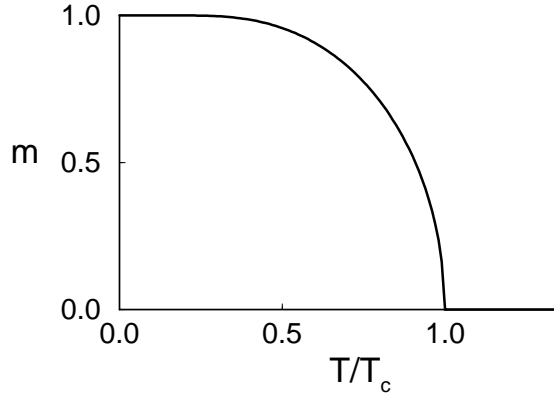


Figure 4.1: Temperature dependence of the order parameter in the mean-field approximation.

The actual local field acting on the site i was replaced by its mean value and is independent of the site, $H_i \rightarrow H_{MF}$. In the *mean-field* approximation the original Hamiltonian is replaced by a mean-field Hamiltonian

$$\mathcal{H}_{MF} = -H_{MF} \sum_i S_i, \quad (4.5)$$

which is equivalent to the Hamiltonian of a non-interacting spin system in an external magnetic field and can be solved exactly:

$$m = \tanh(\beta H_{MF}) = \tanh[\beta(zJm + H)]. \quad (4.6)$$

This is a self-consistent, transcendental equation for m . It has a non-trivial solution ($m \neq 0$) for $\beta zJ > 1$, i.e., at low temperatures. The temperature dependence of the order parameter m for $H = 0$ is shown in Fig. 4.1. It vanishes at the *critical temperature* $T_C = zJ$.

4.1.2 Bragg-Williams Approximation

The Gibbs free energy is a function of T and H ,

$$dG = -SdT - MdH. \quad (4.7)$$

This equation tells us that for fixed T and H , $\delta G = 0$. In equilibrium, thus, G is an extremum of a *trial free energy* \mathcal{G} with respect to any variational parameter. We will now search the trial free energy \mathcal{G} .

We start with the density matrix and calculate the Gibbs free energy of an Ising system. In the Bragg-Williams (mean-field) approach we approximate the total density matrix ρ by a direct product of the density matrices for individual spins,

$$\rho \approx \prod_i \rho_i \propto \prod_i e^{-\beta \hat{\mathcal{H}}_i}. \quad (4.8)$$

Physically this means that the spin fluctuations are considered to be uncorrelated, the orientation of a spin is statistically independent of the orientation of any other spin. There are no correlations in the spin orientation except those coming from the long-range order. Under this assumption, the Ising Hamiltonian is not only diagonal in the spin space (the state $S_i = +1$ is uncoupled from the state $S_i = -1$; physically this means that the Ising system has no dynamics), but it is also independent of all other spins in the system. Thus, the *trial* density matrices ρ_i are diagonal and of the general form (remember, its trace must be = 1):

$$\rho_i = \begin{vmatrix} \frac{1}{2}(1 + m'_i) & 0 \\ 0 & \frac{1}{2}(1 - m'_i) \end{vmatrix}. \quad (4.9)$$

They depend on the variational parameter which is site-independent, $m'_i = m'$, because of translational invariance. With this density matrix the average spin orientation is:

$$\langle S_i \rangle = \text{Tr}(S_i \rho_i) = m. \quad (4.10)$$

Because of statistical independence of the spins in MFA, the short-range spin correlations are:

$$\langle S_i S_j \rangle = \langle S_i \rangle \langle S_j \rangle = m^2. \quad (4.11)$$

We can write the (trial) enthalpy $E = \langle \mathcal{H} \rangle$ [where \mathcal{H} is given by the Eq. (4.1)] as:

$$E = -\frac{1}{2} z J N m'^2 - N H m' \quad (4.12)$$

and the trial entropy as:

$$S = -k_B \text{Tr}(\rho \ln \rho) = -N k_B \left(\frac{1 + m'}{2} \ln \frac{1 + m'}{2} + \frac{1 - m'}{2} \ln \frac{1 - m'}{2} \right)$$

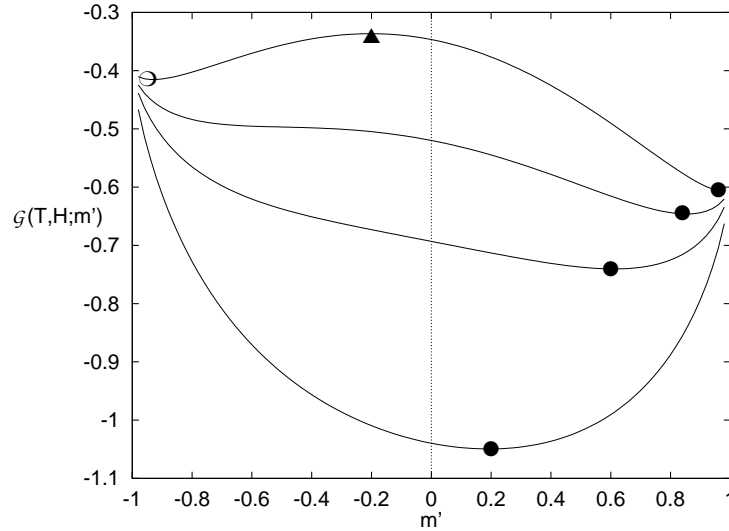


Figure 4.2: Trial Gibbs free energy $\mathcal{G}(T, H; m')$ vs. m' at $H = 0.1 zJ$. Solid circles denote stable states, open circle a metastable state, and the solid triangle an unstable state. The curves are labelled by T/T_C . m and $G(T, H)$ are determined by the minimum of \mathcal{G} .

(4.13)

With these expressions, the trial Gibbs free energy becomes:

$$\mathcal{G}(H, T; m') = E - TS = E + k_B T \text{Tr}(\rho \ln \rho) = -\frac{1}{2} z J N m'^2 - N H m' + N k_B T \left(\frac{1+m'}{2} \ln \frac{1+m'}{2} + \frac{1-m'}{2} \ln \frac{1-m'}{2} \right) \quad (4.14)$$

It depends on H , T , and on the variational parameter m' . \mathcal{G} for $H = 0.1$ (in units of zJ) is shown in Fig. 4.2. To get the true equilibrium Gibbs free energy, $\mathcal{G}(H, T; m)$ has to be *minimized* with respect to the variational parameter m' ,

$$G(T, H) = \min_{m'} \mathcal{G}(H, T; m'). \quad (4.15)$$

The minimization gives:

$$z J m + H = \frac{1}{2} k_B T \ln \frac{1+m}{1-m}, \quad (4.16)$$

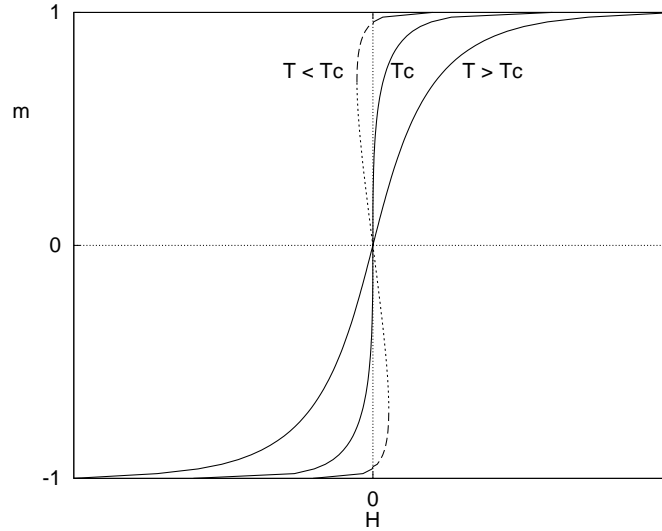


Figure 4.3: External-field dependence of the Ising-model magnetization in the mean-field approximation. T_C is the critical temperature. Solid line: stable states, dashed line: metastable states, dotted line: unstable states.

or:

$$m = \tanh[\beta(zJm + H)]. \quad (4.17)$$

This is the familiar “equation of state” of magnetic systems.

The field-dependence of m , Eq. (4.17) is shown in Fig. 4.3. At high temperatures ($T > T_C$), the order parameter m behaves as in a paramagnet, this is the *paramagnetic phase*. The spins are partially ordered only because of H . For $H = 0$, the spins are *disordered*, therefore this phase is also called the disordered phase. At low temperature ($T < T_C$), the spins are *spontaneously ordered* even in the absence of H . This is the spontaneously ordered phase. The system has *long-range order* for $J > 0$, all the spins are oriented predominantly in the same direction, this is the ferromagnetic phase.

The Gibbs free energy is:

$$\begin{aligned} G(T, H) &= \frac{zN}{2} Jm^2 + \frac{Nk_B T}{2} \ln \frac{1 - m^2}{4} \\ &= \frac{zN}{2} Jm^2 - Nk_B T \ln [2 \cosh \beta(zJm + H)] \end{aligned} \quad (4.18)$$

where m is related to H via the Eq. (4.17). The equation of state is

$$M = - \left(\frac{\partial G(T, H)}{\partial H} \right)_T \quad (M = Nm). \quad (4.19)$$

The Helmholtz free energy is $F(T, M) = G(T, H) + HM$:

$$F_{MF}(T, M) = Nk_B T \left[\frac{1+m}{2} \ln \frac{1+m}{2} + \frac{1-m}{2} \ln \frac{1-m}{2} \right] - \frac{1}{2} N z J m^2. \quad (4.20)$$

The field dependence of the equilibrium Gibbs free energy G is shown in Fig. 4.4. The equation of state now reads

$$H = \left(\frac{\partial F(T, M)}{\partial M} \right)_T. \quad (4.21)$$

The Helmholtz free energy $F(T, M, N)$ is shown in Fig. 4.5.

As before, the critical temperature, where the spontaneous ordering vanishes in the absence of external magnetic field H as T is increased, is determined by the condition $\beta_C z J = 1$,

$$k_B T_C = z |J|. \quad (4.22)$$

In the past, many improvements to the simple MF theory presented here have been proposed. With these improvements, the MF critical temperature came closer to the exact critical temperature. None of these theories, however, was able to bring an improvements to the critical exponents, they kept their mean-field values. The real break-through brought the renormalization-group theory, which will be introduced in Section 4.3.

4.1.3 Landau Theory of Continuous Phase Transitions

In the previous Section we have seen that the trial Gibbs free energy $\mathcal{G}(T, H; M')$ was an analytic function which had approximately a parabolic shape as a function

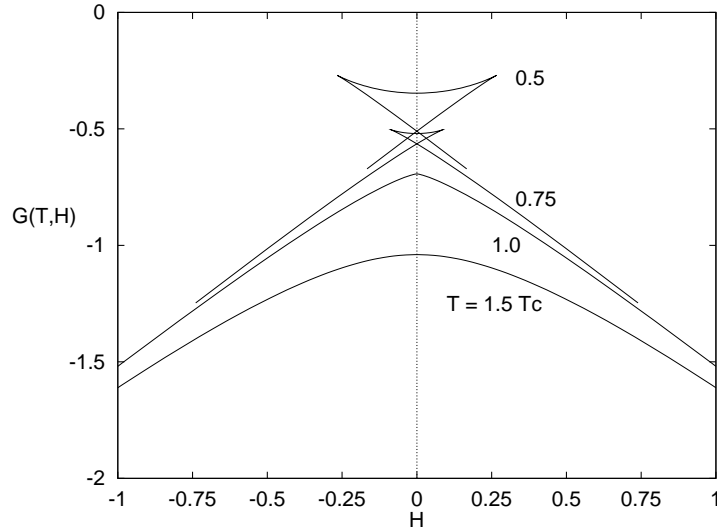


Figure 4.4: Equilibrium Gibbs free energy. At low temperatures, when a first-order transition at $H = 0$ takes place, part of the curve corresponds to metastable states and part to unstable states.

of the variational parameter M' above the critical temperature and a double minimum below T_C . Close to the critical point and for $H = 0$, the minimum of \mathcal{G} is close to $m' = 0$ and we can expand the trial mean-field free energy \mathcal{G} to a very good approximation in a power series of m .

$$\mathcal{G}(T, H; m) = G_0 + \frac{1}{2}rm^2 + um^4 + \dots \quad (4.23)$$

At $T > T_C$, r must be > 0 and below T_C , it must be < 0 , thus r changes sign and in most of the cases it can be approximated by a linear temperature dependence, $r = r'(T - T_C)/T_C$.

The Landau theory is *phenomenological* and deals only with *macroscopic* quantities. The theory does not deal correctly with the fluctuations and is a mean-field theory. In fact, the Landau theory of continuous phase transitions is much more than just simple expansion of \mathcal{G} . It is based on the *symmetry* of the system. Landau postulated that the free energy can be expanded as a power series in m where only those terms that are compatible with the symmetry of the system are allowed.

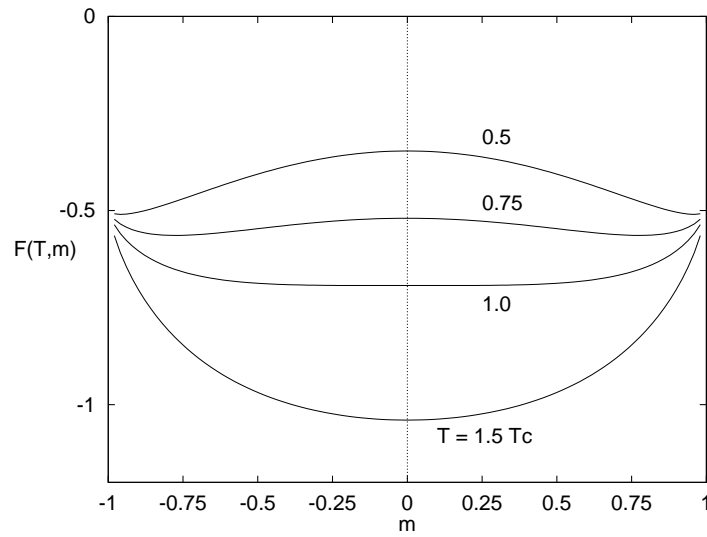


Figure 4.5: Equilibrium Helmholtz free energy. At low temperatures, where the transition is first order, F has two minima, corresponding to $\pm m$. The region between the two minima is the coexistence region when the system can be partially in the $+m$, and partially in the $-m$ state (domain structure!).

Formulation of the Landau Theory

Critical anomalies take place on long length scales. Therefore it is enough and in order to change from discrete to continuous description:

$$\begin{aligned} m_i &\rightarrow m(\vec{r}) \\ h_i &\rightarrow h(\vec{r}). \end{aligned}$$

In formulating the Landau theory let us consider a macroscopic system and assume that there exists a (generalized; not necessarily magnetic) order parameter density $m(\vec{r})$ averaged over *microscopic, atomic* distances a . Because of the average, $m(\vec{r})$ is now a *continuous* function of \vec{r} (continuous in m and in \vec{r})! There can be no fluctuations in $m(\vec{r})$ on the length scale smaller than a ($a \gtrsim$ lattice spacing). Correspondingly, $m(\vec{r})$ has no Fourier components with wavevector greater than $k = 2\pi/a$.

In lattice models, the canonical partition function appropriate to the Gibbs free energy of a spin system is:

$$Y(T, H) = \sum_{\{S_i\}} e^{-\beta \mathcal{H}} \quad (4.24)$$

(we have to sum over all spin configurations). Now, since $m(r)$ is a continuous function of the coordinate r , we must replace the sum by a functional integral over all *functions* $m(r)$. When doing this replacement one must not forget that $m(r)$ is equal to the average of a microscopic spin S_i over a small volume. That means that there can be several microscopic spin configurations which average into the same $m(r)$. Let there be $\Omega(m(r))$ spin configurations that lead to the same $m(r)$. Then, the sum over all spin configurations is equal to the functional integral over all functions $m(r)$ where each function $m(r)$ is weighted by $\Omega(m(r))$:

$$\sum_{\{S_i\}} \rightarrow \int Dm(r) \Omega(m(r)). \quad (4.25)$$

The partition function is:

$$Y(T, H) = \int Dm(r) \Omega(m(r)) e^{-\beta \mathcal{H}(H; m(r))}. \quad (4.26)$$

Using the Boltzmann formula $S = k_B \ln \Omega$, we can write Y in the form:

$$Y(T, H) = \int Dm(r) e^{-\beta \mathcal{H}(H; m(r)) - TS}$$

$$= \int Dm(r) e^{-\beta \mathcal{G}(T, H; m(r))}, \quad (4.27)$$

where we have introduced the *Landau free energy*:

$$\mathcal{G}(T, H; m(r)) = \mathcal{H}(H; m(r)) - TS(m(r)). \quad (4.28)$$

Notice the analogy between this equation and the trial Gibbs free energy obtained previously with the Bragg-Williams method! The Landau free energy is in fact a generalization of the trial Gibbs free energy.

Notice also that the entropy $S(m(r))$ is not equal to the usual microcanonical entropy S (related to the number of configurations in an energy interval ΔE) because we did not make any summation over the energies! $S(m(r))$ is related to the number of microscopic configurations leading to the same function $m(r)$. Only after integration over all the states in a certain energy interval we would get the usual thermodynamic entropy S . Therefore we can call $S(m(r))$ a partial entropy.

We still have to work out the functional integral in order to get the (equilibrium) partition function $Y(T, H)$! From $Y(T, H)$ the Gibbs free energy is:

$$G(T, H) = -k_B T \ln Y(T, H) = -k_B T \ln \int Dm e^{-\beta \mathcal{G}(T, H; m)}. \quad (4.29)$$

In the Landau approach we write down the Landau free energy in a most general form as an expansion in terms of the order parameter and its derivatives by taking into account all the symmetry-allowed terms up to a given order:

$$\mathcal{G}(T, H, m) = \int dV \left(\frac{1}{2} r m^2 + u m^4 + \dots - H m + \frac{1}{2} g |\nabla m|^2 + \dots \right). \quad (4.30)$$

The highest-order coefficient in the expansion must always be positive, else the system were unstable. For the Ising model, this form can also be obtained by expanding G , Eq. (4.14), in powers of m . The last term appears if m is site-dependent:

$$S_i S_j \rightarrow m(r) m(r+a) = m(r) m(r) + m(r) a \nabla m(r) + \frac{1}{2} a^2 \nabla^2 m(r) + \dots \quad (4.31)$$

The first term is already included in the first line of (4.30), the second term vanishes in systems with center of inversion (symmetry argument!) and the last term, after partial integration over the whole volume, gives the term proportional to $|\nabla m|^2$. The prefactor g must also be positive always. A negative g would lead to completely unphysical wild fluctuations in m on extremely short lengthscales.

Here we have assumed a one-component order parameter. In general, the order parameter can have several components or be complex. The coefficients are in principle temperature-dependent and depend also on the symmetry of the system. For example, the coefficient of the cubic term vanishes if the system is invariant under the reversal of $m(r)$ when $H = 0$. Usually we assume that u is temperature independent whereas the quadratic term has a linear temperature dependence,

$$r = r' t, \quad t = (T - T_C)/T_C \quad (4.32)$$

In the Landau theory we assume that the fluctuations of the order parameter are small, that the important values of m lie in a narrow range near its equilibrium value that minimizes the Landau free energy. In fact, the maximal contribution to the partition function comes from the term with minimal Landau free energy. In the thermodynamic limit, the probability of any other configuration will vanish rapidly. Therefore we make use of the saddle-point approximation, the functional integral (4.29) is equal to the maximum value of the integrand, multiplied by a constant. For constant field H , the functional form of $m(r)$ that maximizes the integrand (minimizes the Landau free energy), is a constant, determined by the conditions

$$\frac{\partial \mathcal{G}}{\partial m} = 0 \quad \frac{\partial^2 \mathcal{G}}{\partial m^2} \geq 0 \quad (4.33)$$

In the saddle-point approximation we immediately obtain $G(T, H)$ from Eq. (4.29):

$$G(T, H) = G_0 + \min_m \mathcal{G}(T, H; m) = \quad (4.34)$$

In equilibrium, the system is at the bottom of the Landau free energy \mathcal{G} . This determines m and $G(T, H)$. From $G(T, H)$ we get other thermodynamic quantities, as before.

The advantage of the Landau theory is that it allows one to study a macroscopic system without knowing all its microscopic details. It is based on the symmetry properties of the macroscopic system alone.

4.1.4 Critical Exponents in the Landau Theory

We will now calculate the critical exponents of the Landau model. To be more specific, let us investigate the Landau model of magnetic phase transitions, when

$$\mathcal{G} = G_0 + \left(\frac{1}{2} r m'^2 + u m'^4 - m' H \right) V \quad (4.35)$$

In equilibrium, when $\partial \mathcal{G} / \partial m' = 0$, we get the *equation of state*:

$$r m + 4 u m^3 - H = 0 \quad (4.36)$$

- Order-parameter exponent β ($H = 0$). For $H = 0$, and $r > 0$, the equation of state has a single solution, $m = 0$, there is no spontaneous magnetization above T_C , as expected. For $t < 0$, the equation of state has three roots when $H = 0$. The root at $m = 0$ belongs to an unstable state, so we are left with two real roots. The result is:

$$m^2 = -\frac{r}{4u} \Rightarrow \begin{cases} m = \pm \sqrt{\frac{r'}{4u}} t^{1/2} & \beta = \frac{1}{2} \quad t < 0 \\ m = 0 & t > 0 \end{cases} \quad (4.37)$$

- Critical isotherm exponent δ ($T = T_C$). Now, $r = 0$, and we have $4 u m^3 = H$, from which it follows

$$m = \sqrt[3]{\frac{H}{4u}} \Rightarrow \delta = 3 \quad t = 0 \quad (4.38)$$

- Susceptibility exponent γ . The isothermal susceptibility is defined as

$$\chi_T = \left(\frac{\partial m}{\partial H} \right)_T. \quad (4.39)$$

From the equation of state, Eq. (4.36), we get:

$$r \left(\frac{\partial m}{\partial H} \right)_T + 12 u m^2 \left(\frac{\partial m}{\partial H} \right)_T - 1 = 0 \quad (4.40)$$

from which we find:

$$\chi_T = \begin{cases} \frac{1}{r't} \propto t^{-1} & \Rightarrow \gamma = 1 \quad t > 0 \\ \frac{1}{-2r't} \propto |t|^{-1} & \Rightarrow \gamma = 1 \quad t < 0 \end{cases} \quad (4.41)$$

- Specific-heat exponent α ($H = 0$). The specific heat at constant (zero) field is

$$c_H = -T \left(\frac{\partial^2 G}{\partial T^2} \right)_H \quad (4.42)$$

For $t > 0$, $m = 0$, and $G = G_0 = \text{const.}$ (that means that G_0 does not depend on the reduced temperature t), so c_H is zero. For $t < 0$ (and $H = 0$), $m^2 = -r/4u$ and

$$G = G_0 - \frac{r'^2}{16u} t^2 \quad (4.43)$$

It follows:

$$c_H = \begin{cases} 0 & \Rightarrow \alpha = 0 & t > 0 \\ \frac{r'^2 T_C}{8u} & \Rightarrow \alpha = 0 & t < 0 \end{cases} \quad (4.44)$$

Comment 1: We see that the Landau theory gives mean-field (also called "classical") critical exponents.

Comment 2: At $H \neq 0$, there is no phase transition, all the quantities vary smoothly, only the exponent δ is defined.

4.1.5 Short-Range Order and Fluctuations

Although the correlations and fluctuations are neglected in the MFA, we can still get an estimate of the short-range correlations. To obtain the mean-field values of the exponents ν and η , which describe the behaviour of the correlation function, we have to include the gradient term in the Landau free energy. In the framework of the mean-field theory, this leads to the Ornstein-Zernike approximation. To investigate the correlations, we consider the response of $m(\vec{r})$ to a perturbing field acting, say, at the origin, $H(\vec{r}) = \lambda \delta(\vec{r})$ via the "response function", the susceptibility:

$$m(\vec{r}) = \int d^d r' \chi(\vec{r} - \vec{r}') H(\vec{r}'). \quad (4.45)$$

The Fourier transform of the last convolution is simply:

$$m(\vec{q}) = \chi(\vec{q}) H(\vec{q}). \quad (4.46)$$

To the lowest order in m and ∇m , the Landau free energy is:

$$\mathcal{G}(H; m(\vec{r})) = G_0 + \int \mathrm{d}^d r \left[\frac{1}{2} g |\nabla m(\vec{r})|^2 + \frac{1}{2} r m^2(\vec{r}) + u m^4(\vec{r}) - \lambda m(\vec{r}) \delta(\vec{r}) \right] \quad (4.47)$$

We first make Fourier transforms to the d -dimensional momentum space,

$$\begin{aligned} m(\vec{q}) &= \int \mathrm{d}^d r \, e^{-i\vec{q}\cdot\vec{r}} m(\vec{r}) \\ m(\vec{r}) &= \frac{1}{(2\pi)^d} \int \mathrm{d}^d q \, e^{i\vec{q}\cdot\vec{r}} m(\vec{q}) \end{aligned} \quad (4.48)$$

The Fourier transform of ∇m is:

$$\begin{aligned} \nabla m(\vec{r}) &= \frac{1}{(2\pi)^d} \int \mathrm{d}^d q \, (\nabla e^{i\vec{q}\cdot\vec{r}}) m(\vec{q}) \\ &= \frac{1}{(2\pi)^d} \int \mathrm{d}^d q \, i\vec{q} e^{i\vec{q}\cdot\vec{r}} m(\vec{q}) \end{aligned} \quad (4.49)$$

The Fourier transform of a δ -function is a constant,

$$\delta(\vec{r}) = \frac{1}{(2\pi)^d} \int \mathrm{d}^d q \, e^{i\vec{q}\cdot\vec{r}} \quad (4.50)$$

In the momentum space, the Landau free energy is:

$$\mathcal{G}(H; m(\vec{q})) = G_0 + \int \mathrm{d}^d q \left[\frac{1}{2} g q^2 m^2(\vec{q}) + \frac{1}{2} r m^2(\vec{q}) + u m^4(\vec{q}) - \lambda m(\vec{q}) \right]. \quad (4.51)$$

Here, we approximated the Fourier transform of $m^4(\vec{r})$ by $m^4(\vec{q})$!

In the mean-field approximation the modes with different momenta \vec{q} are assumed to be independent of each other. Therefore we can minimize \mathcal{G} for each individual mode separately,

$$\frac{\partial \mathcal{G}}{\partial m(\vec{q})} = 0, \quad (4.52)$$

and the equation of state in momentum space is:

$$(gq^2 + r)m(\vec{q}) + 4um^3(\vec{q}) = \lambda. \quad (4.53)$$

For $t > 0$, when m vanishes in the absence of the field, we can neglect the cubic term, and we get:

$$m(\vec{q}) = \frac{\lambda}{gq^2 + r}, \quad \chi(\vec{q}) = \frac{1}{gq^2 + r}. \quad (4.54)$$

After a (somewhat tedious) Fourier transform back to real space we find in the limit of large \vec{r} :

$$\chi(\vec{r}) \propto m(\vec{r}) \propto r^{2-d} \exp(-r/\xi) \Rightarrow \eta = 0 \quad (4.55)$$

where

$$\xi = \sqrt{\frac{g}{r't}} \quad \Rightarrow \quad \nu = \frac{1}{2}. \quad (4.56)$$

Below T_C , $m \neq 0$, we replace $m(\vec{r})$ by $m + \delta m(\vec{r})$, where m is the spontaneous magnetization and δm the response to the infinitesimal perturbing field λ . We expand in powers of δm and retain only linear-order terms. We find:

$$\delta m(\vec{q}) = \frac{\lambda}{gq^2 - 2r} \quad (4.57)$$

and the same exponents as for $t > 0$.

In the Ornstein-Zernike approximation (random-phase approximation), the modes with different momenta are considered to be independent of one another. The free energy is a sum (integral) of contributions from non-interacting normal modes.

4.1.6 Validity of The Mean-field Theory

The mean-field theories neglect the fluctuations of the neighbouring spins, each spin interacts only with the mean value of its neighbours. So, the mean-field theories can only be valid when fluctuations are irrelevant. We will now estimate the relevance of fluctuations close to the critical point. We shall calculate the free energy of the fluctuations and compare it with the free energy calculated in the MFA.

To the order of magnitude, only those fluctuations can be thermally excited whose free energy is of the order $k_B T$ (or less). The free energy of a typical fluctuation will thus be proportional to $k_B T$. Since the size of the fluctuation is of

the order of the correlation length ξ (this is how the correlation length is defined), the free energy of a typical fluctuation *per unit volume* is proportional to

$$G_{fluct} \propto \xi^{-d} \propto |t|^{\nu d}. \quad (4.58)$$

On the other hand, the singular part of the mean-field free energy is governed by the specific-heat exponent α :

$$G_{MF} \propto |t|^{2-\alpha} \quad (4.59)$$

(we obtain this expression after two integrations of the specific heat over temperature). For the mean-field theory to be "exact", $G_{fluct} \ll G_{MF}$ in the limit as $t \rightarrow 0$,

$$\begin{aligned} \lim_{|t| \rightarrow 0} \frac{G_{fluct}}{G_{MF}} &\propto \lim_{|t| \rightarrow 0} |t|^{\nu d - 2 + \alpha} = 0 \\ \Rightarrow \nu d - 2 + \alpha &> 0 \Rightarrow d > \frac{2 - \alpha}{\nu}. \end{aligned} \quad (4.60)$$

For the Ising model (scalar order parameter) in the mean-field approximation $\alpha = 0$ and $\nu = 1/2$ and we get the condition

$$d > 4. \quad (4.61)$$

The mean-field exponents are correct and the fluctuations are irrelevant only for $d > 4$. $d = 4$ is called the *upper critical dimension* of the Ising model.

4.2 Monte Carlo Simulations

With the modern computers, different computer simulations became extremely powerful techniques. One sets up in the computer the Hamiltonian describing a physical system and tries to see how it behaves. However, an efficient and reliable computer simulation is full of tricks and a lot of experience is needed to produce reliable results.

Among different computer simulations, the Monte Carlo (MC) simulations are probably the most widely used in statistical mechanics. MC simulations are used to study classical as well as quantum systems, continuous spin systems, liquids and other disordered systems, gauge theories, etc. Here we shall explain the principles of the MC methods on the prototype example, the Ising model on a square lattice.

4.2.1 Importance Sampling

In statistical mechanics we calculate the expectation values of thermodynamic observables. For a spin system, the expectation value is defined as

$$\langle A \rangle = \frac{\sum_{\{s\}} A e^{-\beta \mathcal{H}}}{\sum_{\{s\}} e^{-\beta \mathcal{H}}}. \quad (4.62)$$

For an Ising system, the sum is over 2^N configurations. This number is extremely large for all but very small systems. Therefore one must find a method which estimates thermodynamic properties by sampling a small subset of representative configurations. One possible strategy would be to scan randomly the whole configuration space (scan randomly over all spin configurations). The expectation value of A would then be obtained as a sum over the scanned configurations, weighted by the appropriate Boltzmann factors. However, this random sampling has a very serious drawback. Because of the Boltzmann factor which brings a negligible weight to most of the configurations, very few configurations will contribute to the expectation value of A and a very unreliable estimate will result. This problem occurs because only an extremely restricted part of the configuration space is occupied with a considerable probability in the thermodynamic limit. Therefore it makes sense to restrict the sampling only to these states. this is the *importance sampling*.

Importance sampling is realized if we generate a *Markov chain* of configurations, i.e., a configuration is generated from the preceding configuration. The disadvantage of this Markov chain is that the successive configurations are strongly

correlated and therefore very long sequence of configurations is needed to obtain reliable averages. The expectation value of A is now equal to:

$$\langle A \rangle = \lim_{n_{\max} \rightarrow \infty} \left[\frac{1}{n_{\max} - n_{\min}} \sum_{n=n_{\min}}^{n_{\max}} A_n \right] \quad (4.63)$$

where A_n is the average of A over n successive configurations.

4.2.2 Metropolis Algorithm

In generating successive configurations, the transition probability must be normalized, ergodic, and it must obey detailed balance. The Metropolis algorithm fulfills the above conditions. A new configuration $\{S_f\}$ is obtained from the previous configuration $\{S_i\}$ by flipping one or more spins. In the Metropolis algorithm, the probability that the system changes from $\{S_i\}$ to $\{S_f\}$ is

$$P(\{S_i\} \rightarrow \{S_f\}) = \begin{cases} e^{-\beta(E_f - E_i)} & \text{if } E_f > E_i \\ 1 & \text{if } E_f \leq E_i \end{cases} \quad (4.64)$$

where E_i and E_f are the energies of the previous and final states, respectively. With this choice of transition probabilities, the system tends asymptotically to a steady state in which the probability of a given configuration is $e^{-\beta E(\{S\})}$.

The practical implementation of the Metropolis algorithm is very simple. First, a spin is selected either at random or sequentially. The spin is flipped and the energies E_i and E_f are calculated. Finally, a pseudo-random number is used to accept or reject the flip with the probability (4.64). The meaning of this condition is straightforward. If the system gains in energy by flipping the spin, the flip is accepted. It is also accepted if it costs energy to flip the spin, but the energy difference is small so that the Boltzmann factor is larger than the random number (analogy to thermal excitations!). The flow chart of a MC program is shown in Fig. 4.6.

4.2.3 Practical Details and Data Analysis

- *Initial Configuration.* In principle, the initial configuration should not influence the final configuration. However, the equilibrium may be reached faster if a proper initial configuration is chosen. An example, showing the approach to thermal equilibrium in a $d = 3$ Ising ferromagnet, is shown in Fig. 4.7.

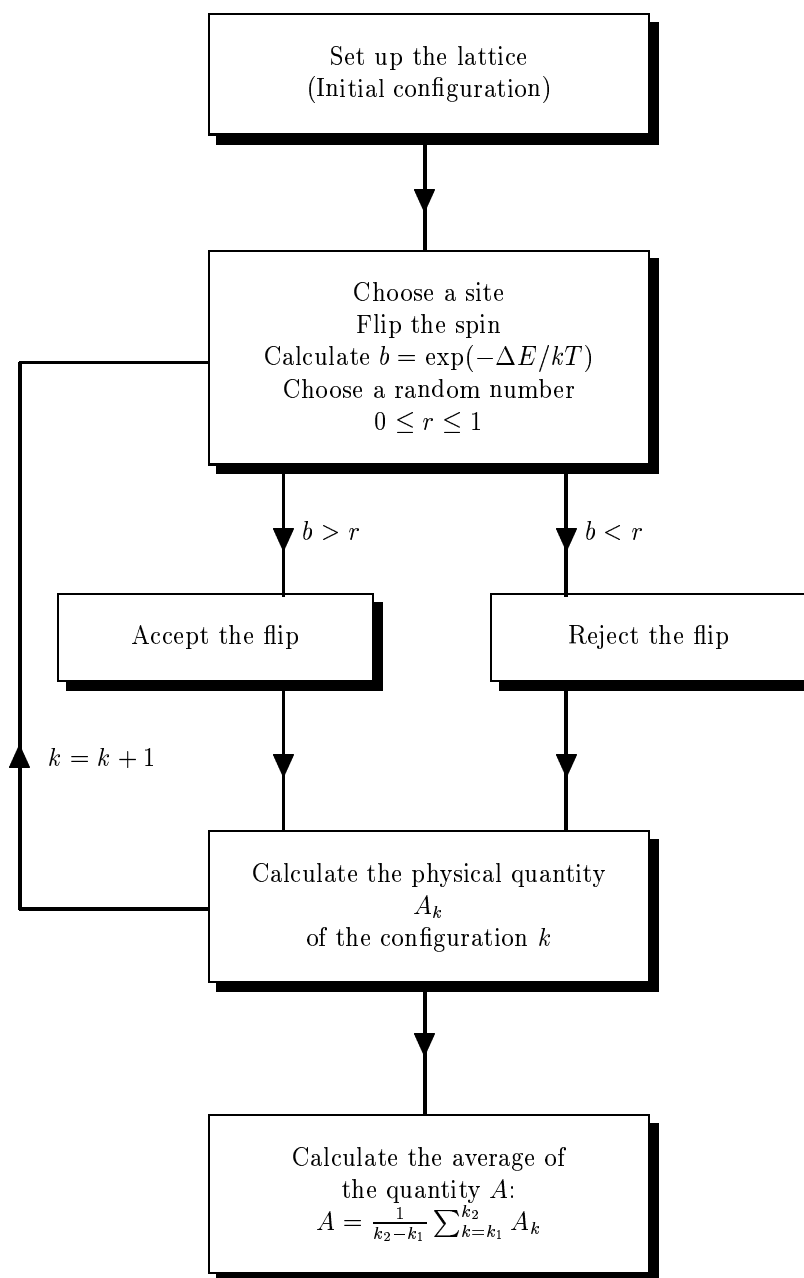


Figure 4.6: Flow diagram of a Monte-Carlo simulation with Metropolis algorithm. During the first k_1 steps the system still approaches to its equilibrium configuration, therefore these configurations are not included in calculating the average properties.

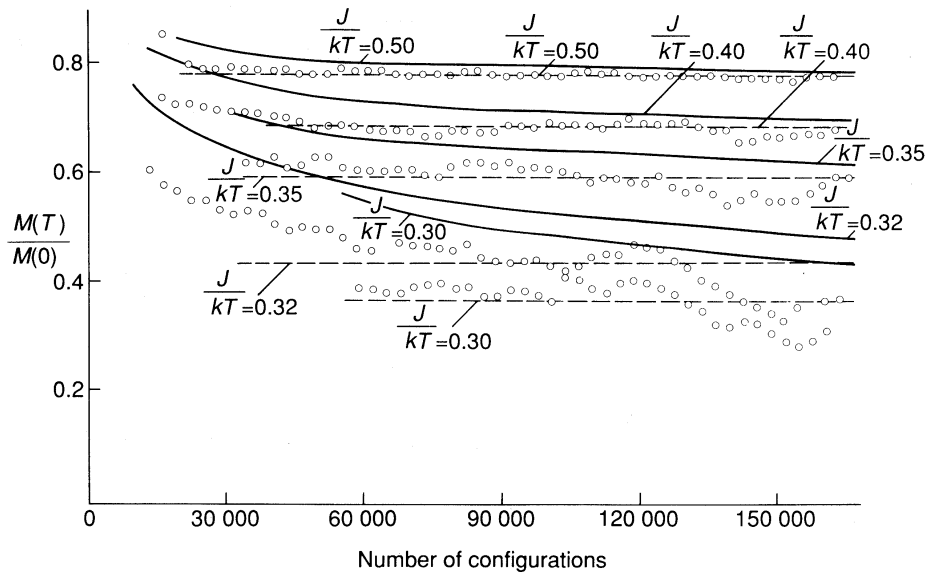


Figure 4.7: Relaxation of magnetization towards its equilibrium value in a Monte-Carlo simulation of an Ising system on a $10 \times 10 \times 10$ lattice with periodic boundary conditions. Solid lines are “instantaneous” magnetizations and dashed lines their equilibrium estimates. (From: K. Binder and Z. Rauch, *Z. Phys.* **219**, 201 (1969).)

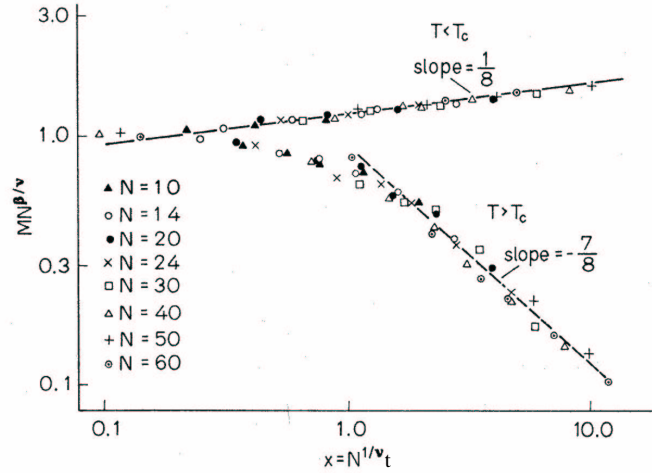


Figure 4.8: Scaled magnetization vs. scaled temperature of an Ising model on a square lattice. (From: D.P. Landau, Phys. Rev. **B 13**, 2997 (1976).)

- *Boundary Conditions.* Different boundary conditions are used. In periodic boundary conditions the system is periodically repeated in all directions to mimic an infinite system. Alternatively, one may describe the effect of the surrounding spins on an (inevitably) finite lattice by a mean field, acting on the spins at the border. Open boundary conditions, when the finite lattice is isolated and the spins at the surface have less neighbours than other spins, have also been used.
- *Finite-size effects.* Lattices, considered in MC simulations, must have a finite size (L) whereas we are usually interested in the macroscopic properties. The question is, how to take the limit of L to ∞ . Usually, one calculates an average for several different lattice sizes and then extrapolates the average to infinite L . This extrapolation is simple away from the critical point where the correlation length is small compared to L . The problem becomes non-trivial in the critical region because ξ which would normally diverge, cannot exceed L . As a consequence, singularities associated with the critical point are rounded. If possible, scaling forms should be used in extrapolating $L \rightarrow \infty$ (see Fig. 4.8).
- *Random-number generator.* Generation of uncorrelated random numbers is

not a trivial problem. Computer languages have built-in "pseudo-random-number" generators. Not always these pseudo-random-number generators are random. In particular, if random sampling is used, then it might happen that some points on the lattice are visited more often than the other (which - in the worst case - may never be visited). Pseudo-random-number generators use algorithms to calculate a sequence of (random) numbers. Care must be taken to initialize random number generators with different "seeds" in order not to obtain the same sequence of random numbers each time.

- *Statistical errors.* To obtain reliable results, an average must be taken over n much larger than the number over which the MC states are correlated. This becomes more difficult near the critical point because of critical slowing down.
- *Free energy.* Monte-Carlo simulation samples only a very small part of the phase space whereas to obtain the partition function or the free energy, we must sum over the whole configuration space. In MC simulations, therefore we calculate the free energy by integrating one of its derivatives, e.g., $-\int M dH$. This means that we need to calculate $M(H)$ in a range of H .
- *Conservation laws.* Above, the spins were allowed to flip uncorrelated from one another. This works fine for magnetic systems, where the order parameter is not conserved. For binary alloys, e.g., however, the number of atoms is conserved and so is the order parameter. So, opposite spins have to flip simultaneously (one "up" \rightarrow "down" and another - usually a neighbouring - one "down" \rightarrow "up"). The conservation laws have important consequences for the kinetics of fluctuations (slowing down).

4.3 The Renormalization Group Methods

Some typical results of MC simulations of a $d = 2$ Ising ferromagnet are shown in Fig. 4.9. At high T ($T = 2T_C$), there is only short-range order, the spins form small clusters. The correlation length (approximately equal to the linear size of the largest cluster) is small. Close (but above) T_C , somewhat larger patches in which most of the spins are lined up in the same direction begin to develop. When the system reaches T_C , these patches expand to infinite size, but fluctuations of smaller scale persist! At the critical temperature, spins form clusters at all lengthscales, including one infinite-size cluster (which cannot be seen on a finite system, but we know that the correlation length has to diverge at T_C). As a result, *all scales of length must be included in a theoretical description*. Another very important conclusion is reached if we substitute a block of 3×3 spins by a single spin with the orientation prescribed by the majority rule (block spin "up" if the majority of the spins in the block is "up", and vice versa). In this way the length scale of the lattice is changed by a factor $b = 3$ each time. This is a "real-space block-spin renormalization-group transformation".

Examples of such transformations are shown in Figs. 4.10-4.12. Below T_C , the system has long-range order but, because the system is close to T_C , the fluctuations form large clusters of oppositely ordered spins. However, under each transformation the size of the unit cell increases, the fluctuations become smaller and smaller and those fluctuations, which are smaller in size than the lattice spacing, disappear. Under successive transformations, thus, the system becomes more and more ordered, it effectively approaches zero temperature, see Fig. 4.13.

Above T_C (Fig. 4.11), there is no long-range order, spins form random "up" and "down" clusters. Under each transformation, the correlation length (expressed in units of the new unit cell) decreases, the clusters become smaller and smaller as if the temperature were higher, but they never disappear. At $T = T_C$ (Fig 4.12), the clusters are of *all length scales*, including one infinitely large cluster. The correlation length is infinite (exceeds the size of the lattice used in the MC simulation). Under successive transformations, it remains "infinite", there is no change in the distribution of cluster sizes, the system is invariant under scale transformations! The system remains at the critical temperature and the critical point is a *fixed point*. Other (trivial) fixed points are $T = 0$ and $T = \infty$ because the system away from T_C tends to one of these points under successive scale transformations. Decrease of the correlation length under successive transformations is an extremely useful property because it makes the fluctuations uncorrelated and we can solve the system using an approximate (e.g., mean-field) theory to calculate

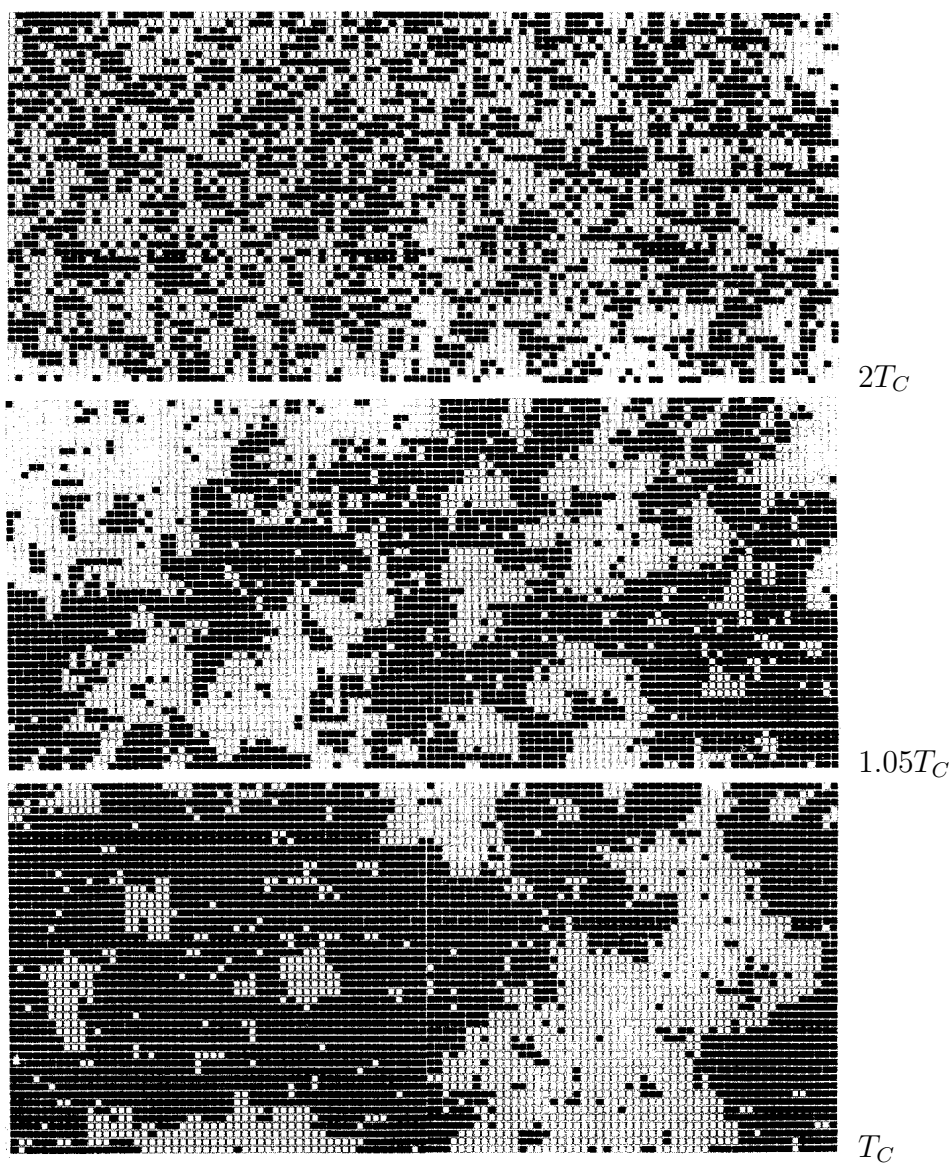


Figure 4.9: Results of MC simulations of a $d = 2$ Ising ferromagnet at different T . At high T ($T = 2T_C$; top), the correlation length is small. Close (but above) T_C (middle), somewhat larger clusters begin to develop. When the system reaches T_C (bottom), these clusters expand to infinite size, but fluctuations on smaller scale persist! (From: K.G. Wilson, Sci. Am. **241**, 140 (1979).))

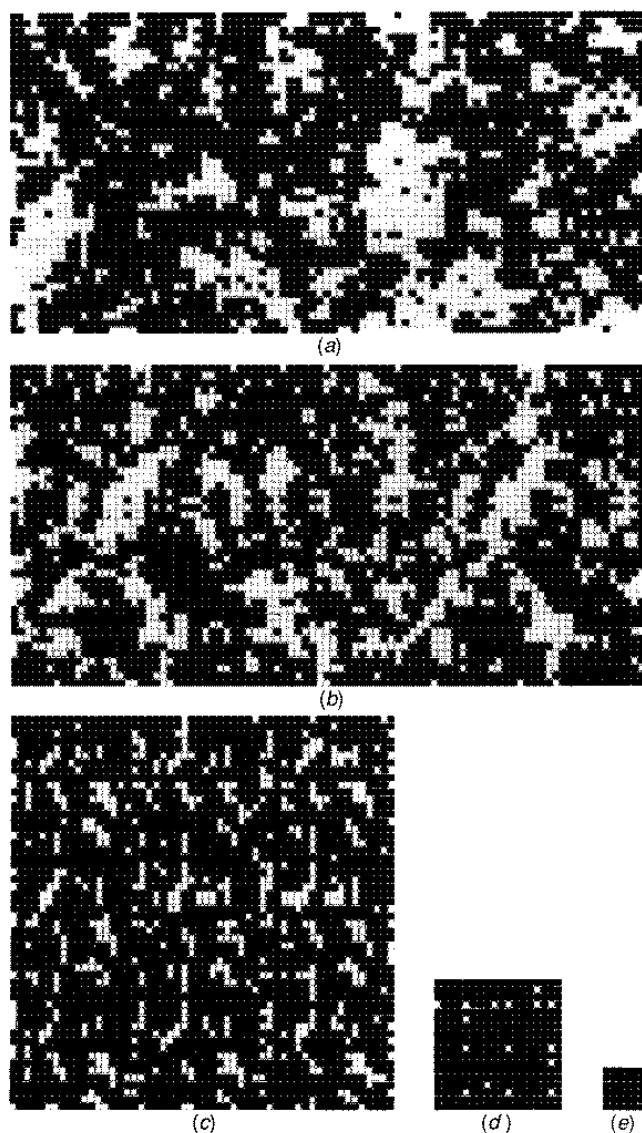


Figure 4.10: Block-spin transformation is applied to a lattice repeatedly at $T = 0.99T_C$. Each time the number of spins is reduced by a factor of 9, elucidating the behaviour of the system at a larger scale. The correlation length decreases under successive transformations. Clearly seen is the decrease of the correlation length and a suppression of the fluctuations. Under each transformation the system appears more ordered, the system flows under the renormalization transformations towards zero temperature. (After K.G. Wilson, Sci. Am. **241**, 140 (1979))

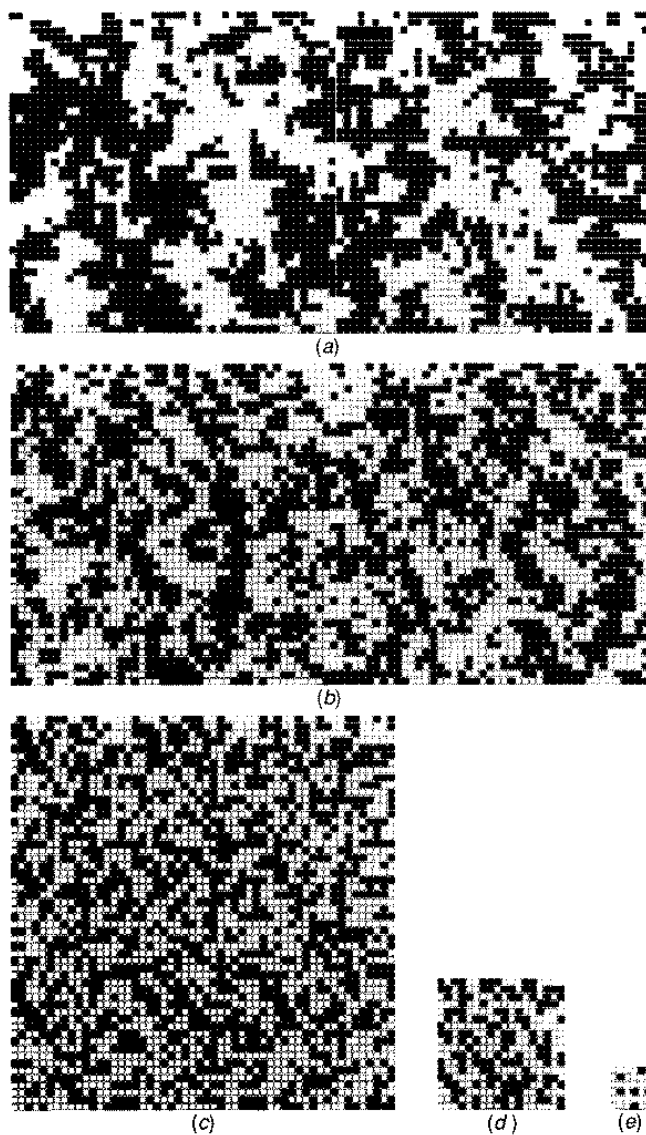


Figure 4.11: As Fig. 5.9 but for $T = 1.22T_C$. Under each transformation, the correlation length decreases, the system flows towards infinite temperature. (After K.G. Wilson, Sci. Am. **241**, 140 (1979))

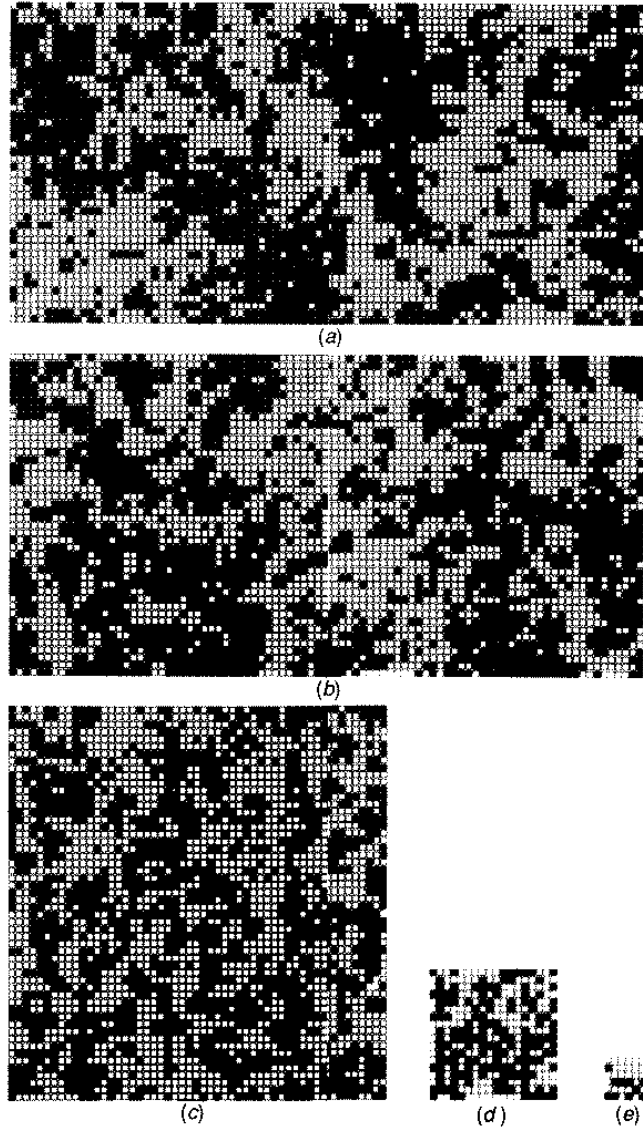


Figure 4.12: As Fig. 5.9 but for $T = T_C$: The correlation length is infinite and remains infinite under successive transformations, the system remains at the critical temperature. (After K.G. Wilson, Sci. Am. **241**, 140 (1979))

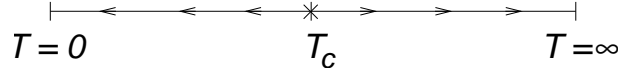


Figure 4.13: Block-spin renormalization at $H = 0$. Upon successive decimations of an initial configuration at $T < T_C$ the system becomes more and more ordered as if the temperature were lower and lower. Eventually, the system becomes completely ordered, this is the $T = 0$ fixed point. Upon decimations of a state at $T > T_C$, the system becomes more and more disordered, the correlation length decreases as if the temperature were higher, until the correlation length becomes on the order of the lattice spacing. This is the infinite-temperature fixed point. A system at T_C stays, upon successive decimations, at T_C . This is another fixed point, the critical point.

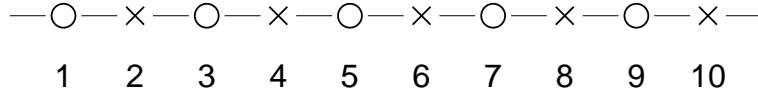
the properties in that region and then transform back to the original lattice. This is the basic idea of the renormalization-group methods and we shall exploit this idea on some simple examples.

4.3.1 A Trivial Example: The $d = 1$ Ising Model

In the renormalization-group methods the temperature changes under successive transformations, therefore it is convenient to work with the *reduced Hamiltonian*, we divide \mathcal{H} by $k_B T$. The reduced Hamiltonian of an Ising Hamiltonian with NN exchange interaction in an external field is

$$\mathcal{H} = -K \sum_i S_i S_{i+1} - h \sum_i S_i, \quad K = \frac{J}{k_B T} \quad h = \frac{H}{k_B T}. \quad (4.65)$$

Recursion relations. If we want to make several successive renormalization-group (RG) transformations, the structure of \mathcal{H} must not change after each transformation. We are thus seeking for such a transformation from the old Hamiltonian \mathcal{H} to the new Hamiltonian \mathcal{H}' that keeps the structure of the Hamiltonian unchanged under successive transformations. In addition, the transformation must be such that it *does not* change the partition function. That is necessary because this ensures that the physical properties of the system are not changed. The partition function of the one-dimensional Ising model in an external field is written in

Figure 4.14: Decimation on the $d = 1$ Ising model.

the form

$$\begin{aligned}
 Z(N, K, h) &= \sum_{\{S\}} e^{-\mathcal{H}} = \sum_{\{S\}} e^{K \sum_i S_i S_{i+1} + h \sum_i S_i} \\
 &= \sum_{\{S\}} e^{K(S_1 S_2 + S_2 S_3) + h S_2 + \frac{h}{2}(S_1 + S_3)} \\
 &\quad \times e^{K(S_3 S_4 + S_4 S_5) + h S_4 + \frac{h}{2}(S_3 + S_5)} \times \dots, \quad (4.66)
 \end{aligned}$$

We divide the lattice into the sublattices \times and \circ , denote the spins on the \times sublattice by σ_i (Fig. 4.14),

$$Z = \sum_{\{S\}} \sum_{\{\sigma\}} e^{\sigma_2 [K(S_1 + S_3) + h] + \frac{h}{2}(S_1 + S_3)} e^{\sigma_4 [K(S_3 + S_5) + h] + \frac{h}{2}(S_3 + S_5)} \dots, \quad (4.67)$$

and carry out a partial trace, i.e., we sum over $2^{N/2}$ configurations of the $\{\sigma\}$ spins sitting on the \times sublattice:

$$Z(N, K, h) = \sum_{\{S\}} \left[e^{(K + \frac{h}{2})(S_1 + S_3) + h} + e^{(-K + \frac{h}{2})(S_1 + S_3) - h} \right] \times \dots. \quad (4.68)$$

The entire partition function must not change and \mathcal{H} must keep its structure, therefore we have the condition

$$Z(N, K, h) = e^{Ng(K, h)} Z\left(\frac{N}{2}, K', h'\right) = e^{Ng} \sum_{\{S\}} e^{-\mathcal{H}'} \quad (4.69)$$

where

$$\mathcal{H}' = -K' \sum_{i \in \text{odd}} S_i S_{i+2} - h' \sum_{i \in \text{odd}} S_i \quad (4.70)$$

and g is a constant. This will ensure that the physical properties will not change under successive iterations. The two conditions (4.70) and (4.69) imply that

$$e^{(K + \frac{h}{2})(S_i + S_{i+2}) + h} + e^{(-K + \frac{h}{2})(S_i + S_{i+2}) - h}$$

$$= e^{K' S_i S_{i+2} + \frac{h'}{2} (S_i + S_{i+2}) + 2g}. \quad (4.71)$$

The last equation has to hold for any $S_i, S_{i+2} = \pm 1$:

$$\begin{aligned} e^{2K+2h} + e^{-2K} &= e^{K'+h'+2g} \\ e^{2K-2h} + e^{-2K} &= e^{K'-h'+2g} \\ e^{+h} + e^{-h} &= e^{-K'+2g}, \end{aligned} \quad (4.72)$$

from which we get:

$$\begin{aligned} e^{4K'} &= \frac{(e^{2K+2h} + e^{-2K})(e^{2K-2h} + e^{-2K})}{(e^h + e^{-h})^2} \\ e^{2h'} &= \frac{e^{2K+2h} + e^{-2K}}{e^{2K-2h} + e^{-2K}} \\ e^{8g} &= (e^{2K+2h} + e^{-2K})(e^{2K-2h} + e^{-2K})(e^h + e^{-h})^2 \end{aligned} \quad (4.73)$$

or:

$$\begin{aligned} K' &= \frac{1}{4} \ln \frac{\cosh(2K+h) \cosh(2K-h)}{\cosh^2 h} \\ h' &= h + \frac{1}{2} \ln \frac{\cosh(2K+h)}{\cosh(2K-h)} \\ g &= \frac{1}{8} \ln [16 \cosh(2K+h) \cosh(2K-h) \cosh^2 h]. \end{aligned} \quad (4.74)$$

Equations (4.73) or (4.74) are the *recursion relations* and determine the fixed points and the flow diagram of the system.

In each iteration, the number of degrees of freedom is reduced by one half, the new Hamiltonian \mathcal{H} has only one half of the previous spins,

$$N' = N/b \quad b = 2, \quad (4.75)$$

and the lattice spacing is increased,

$$a' = ba. \quad (4.76)$$

Other lengths which are measured in units of the lattice spacing are reduced by a factor b . For example, the new correlation length is

$$\xi' = \xi/b. \quad (4.77)$$

The remaining spins on the decimated lattice interact with their new nearest neighbours through the renormalized coupling constants K' and are subject to renormalized fields h' .

Fixed points. At a fixed point the parameters (the coupling constants – in our case K and h) do not change under successive decimations. This means that the system at a fixed point stays at this point in the parameter space. With the new variables

$$\begin{aligned} x &= e^{-4K} & y &= e^{-2h} & z &= e^{-8g} \\ & & & & 0 \leq (x, y, z) &\leq 1 \end{aligned} \quad (4.78)$$

the recursion relations become

$$\begin{aligned} x' &= x \frac{(1+y)^2}{(x+y)(1+xy)} \\ y' &= y \frac{x+y}{1+xy} \\ z' &= z^2 xy^2 \frac{1}{(x+y)(1+xy)(1+y)^2} \end{aligned} \quad (4.79)$$

The recursion relations for x and y do not depend on z . Physically, this means that the singular behaviour of the free energy does not depend on a shift in the energy scale. Therefore, we first investigate the flow diagram and fixed points in the (x, y) plane, see Fig. 4.15. From the first equation it is easy to see that $x' = x$ if $x = 1$, i.e., $x=1$ is a fixed point of this system for arbitrary y . We denote the fixed points by a $*$. There is a line of fixed points at $x^* = 1$ for $0 \leq y^* \leq 1$, corresponding to an infinite-temperature sink. Along this line, the spins are completely disordered and $\xi \rightarrow 0$. If we set $x = 0$, the second equation reduces to $y' = y^2$, which has two fixed points, one at $y^* = 1$ and one at $y^* = 0$. The fixed point at $(x^*, y^*) = (0, 1)$ (this means $H = 0$ and $T = 0$) corresponds to the critical point where $\xi \rightarrow \infty$. This is correct because we know that the $d = 1$ Ising model has $T_C = 0$. This fixed point is unstable and any point in its vicinity will flow under RG transformations away from it towards the fixed points on the line at $x^* = 1$. The critical behaviour of the $d = 1$ Ising model is governed by the least stable fixed point at $(x^*, y^*) = (0, 1)$. The point at $(x^*, y^*) = (0, 0)$ is another fixed point which corresponds to $T = 0$ and $H = \infty$. At this fixed point, the spins are completely ordered.

4.3.2 RG Transformations on the $d = 2$ Ising Model

The recursion relations of the $d = 1$ Ising model could be formulated exactly. Usually, however, farther-range interactions are introduced under successive dec-

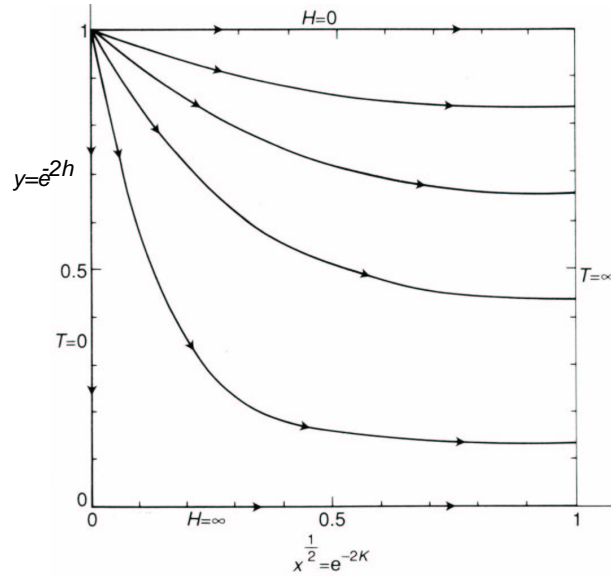


Figure 4.15: Schematic flow diagram of the $d = 1$ Ising model. (After J.M. Yeomans, *Statistical Mechanics of Phase Transitions*, Oxford 1992.)

imations and approximations have to be made in order to get the closed-form recursion relations.

As an example, let us apply the decimation on the $d = 2$ Ising model on a square lattice in the absence of an external field. We start with NN (nearest neighbours) exchange interactions and sum the partition function over all possible configurations of the σ spins (these are the spins that will be decimated, they are denoted by \times in Fig. 4.16).

$$\begin{aligned}
 Z &= \sum_{\{S\}} e^{K \sum_{\langle ij \rangle} S_i S_j} = \sum_{\{S\}} \sum_{\{\sigma\}} e^{K \sum_i (S_{i+x} + S_{i-x} + S_{i+y} + S_{i-y}) \sigma_i} \\
 &= \sum_{\{S\}} \prod_i \left[e^{K(S_{i+x} + S_{i-x} + S_{i+y} + S_{i-y})} + e^{-K(S_{i+x} + S_{i-x} + S_{i+y} + S_{i-y})} \right] \dots
 \end{aligned}
 \tag{4.80}$$

$S_{\pm x}$ and $S_{\pm y}$ are the "right/left" and "upper/lower" neighbours of σ_i , respectively. The conditions are that under decimation Z must not change and that \mathcal{H} should have the same structure as the original Hamiltonian. However, we soon realize that we cannot fulfill the above conditions with the renormalized form of the initial

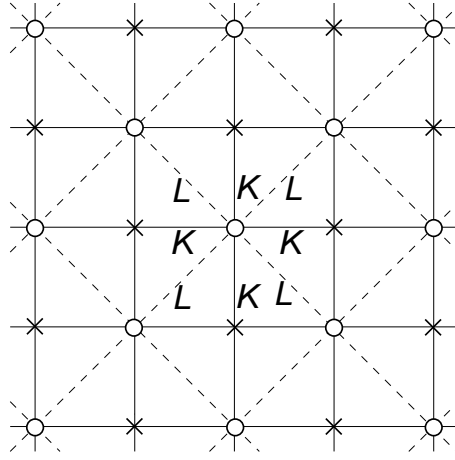


Figure 4.16: Decimation on the $d = 2$ square lattice. The spins denoted by \times are decimated. K is the NN interaction and L the NNN interaction. After decimation, the lattice constant increases by $b = \sqrt{2}$, and previous NNN become NN (nearest neighbours). Solid lines represent the original lattice and dashed lines the decimated lattice.

Hamiltonian. After decimation we get a more complicated Hamiltonian:

$$\begin{aligned} \mathcal{H}' = & - \sum_i [K'(S_{i+x}S_{i+y} + S_{i+y}S_{i-x} + S_{i-x}S_{i-y} + S_{i-y}S_{i+x}) \\ & + L'(S_{i+x}S_{i-x} + S_{i+y}S_{i-y}) + Q'S_{i+x}S_{i+y}S_{i-x}S_{i-y} + C'], \end{aligned} \quad (4.81)$$

where K' is the new NN interaction, L' the new NNN (next nearest neighbours) interaction, Q' the four-spin interaction, and C' a constant. Only with the Hamiltonian of this form, it is possible to satisfy the condition $Z' = Z$ for arbitrary configurations of $S_{\pm x, y}$. Although we started with only NN interactions, decimation generated NNN interaction and even four-spin interactions. So, we have to start from the beginning again with a more complicated Hamiltonian which includes all these interactions. In the following we shall neglect Q' and assume that L is small. In this case we obtain the following approximate recursion relations:

$$\begin{aligned} K' & \approx L + 2K^2 \\ L' & \approx K^2 \end{aligned} \quad (4.82)$$

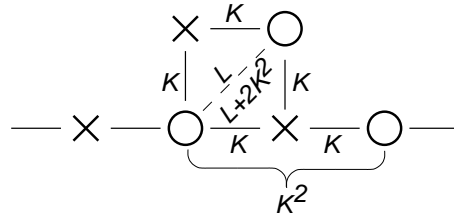


Figure 4.17: The renormalized interaction after a decimation is $L + 2K^2$.

In fact, these relations could be guessed immediately, see Fig. 4.17. The new nearest neighbours are the previous NNN, therefore the term L in the first equation. Besides, they were connected by two pairs of consecutive bonds, this brings $2K^2$. In the second equation, the new next nearest neighbours (NNN) are connected by two consecutive NN bonds on the old lattice - each of them contributing a factor K .

The recursion relations 4.82 have two trivial fixed points, $(K^*, L^*) = (0, 0)$ and $(K^*, L^*) = (\infty, \infty)$ and one non-trivial fixed point at $(1/3, 1/9)$. The first one is the infinite-temperature, paramagnetic fixed point, the second one is the zero-temperature, ferromagnetic fixed point. The non-trivial fixed point corresponds to the critical point of the system.

4.3.3 General Properties of the RG Transformations

So far, we have seen that the renormalization-group transformations work by changing the length scale and by reducing the number of the degrees of freedom. Only at criticality the system remains unaltered under the change in scale and the critical behaviour is determined by a fixed point of the transformation.

We start with an effective Hamiltonian, which is characterized by a set of coupling constants and fields K, L, \dots . Each set of these constants defines a point in the "Hamiltonian space," the space of coupling constants. The renormalization-group transformations work in this space,

$$\mathcal{H}' = \mathcal{R} \mathcal{H}. \quad (4.83)$$

The renormalization-group operator \mathcal{R} reduces the number of degrees of freedom

from N to N' . This defines the scale factor b :

$$b^d = \frac{N}{N'}. \quad (4.84)$$

In the previous examples, the reduction was made in real space either by block-spin transformation or by decimation; later we shall do the reduction also by integrating out large momenta (short wavelengths) in momentum space.

The essential condition to be satisfied by any RG transformation is that the partition function must not change,

$$Z_{N'}(\mathcal{H}') = Z_N(\mathcal{H}). \quad (4.85)$$

Therefore the total free energy does not change but the free energy per unit cell (spin) increases as

$$f(\mathcal{H}') = b^d f(\mathcal{H}). \quad (4.86)$$

All lengths, measured in units of the new lattice spacing, are reduced by the factor b . Thus, the correlation length scales as

$$\xi' = \frac{1}{b} \xi. \quad (4.87)$$

Fixed points. The iteration of \mathcal{R} traces a trajectory in the Hamiltonian space, this is called the "Hamiltonian flow." The trajectories end in fixed points \mathcal{H}^* , defined by

$$\mathcal{H}^* = \mathcal{R}\mathcal{H}^* \quad \Longleftrightarrow \quad \mathcal{H}' = \mathcal{H} \equiv \mathcal{H}^*. \quad (4.88)$$

At the fixed points, the system is invariant under subsequent scale changes and thus the correlation length does not change

$$\xi' = \xi \equiv \xi^*. \quad (4.89)$$

Let us now take a closer look at the Hamiltonian flow near the fixed points. We introduce a reduced Hamiltonian which we write in the general form

$$\mathcal{H} = \vec{\mu} \cdot \vec{\phi} \quad (4.90)$$

and investigate the flow in the *parameter space* spanned by the generalized fields $\vec{\mu}$, conjugate to the operator $\vec{\phi}$ which contains product of operators (like spins).

- *Example:* In the case of the $d = 2$ Ising model on a square lattice, the components of $\vec{\phi}$ include the terms linear in S (the conjugate field is h), quadratic in S (coming from interactions with NN, NNN neighbours), four-spin terms, etc. The corresponding conjugate fields are the magnetic field h and the coupling constants K, L , etc. In the beginning, many components of $\vec{\mu}$ are zero, but under successive transformations, more and more complex interactions emerge and an increasing number of components of $\vec{\mu}$ becomes $\neq 0$.

Under successive RG transformations, the Hamiltonian evolves and the system moves through the parameter space:

$$\vec{\mu}' = \mathcal{R}\vec{\mu} \quad (4.91)$$

until it reaches a fixed point when:

$$\vec{\mu}' = \vec{\mu} \equiv \vec{\mu}^*. \quad (4.92)$$

Every particular type of criticality corresponds to a particular fixed point and deviations from the critical point are related to deviations from the corresponding fixed point. To get insight into the critical properties we must investigate the flow in the Hamiltonian space in the *vicinity* of the fixed points.

Near a fixed point we linearize the operator \mathcal{R} . We write

$$\begin{aligned} \vec{\mu} &= \vec{\mu}^* + \delta\vec{\mu} \\ \vec{\mu}' &= \vec{\mu}^* + \delta\vec{\mu}' \end{aligned} \quad (4.93)$$

where the small deviations from $\vec{\mu}^*$ are related by

$$\vec{\mu}' = \mathcal{R}(\vec{\mu}^*) + \left. \frac{\partial \mathcal{R}}{\partial \vec{\mu}} \right|_{\vec{\mu}=\vec{\mu}^*} \delta\vec{\mu} = \vec{\mu}^* + \mathcal{L}(\vec{\mu}^*) \delta\vec{\mu} \quad (4.94)$$

$$\delta\vec{\mu}' = \mathcal{L}(\vec{\mu}^*) \delta\vec{\mu}. \quad (4.95)$$

The linear matrix \mathcal{L} is evaluated at the fixed point and is thus a constant. The eigenvalues λ_i and the eigenvectors \vec{v}_i of \mathcal{L} ,

$$\mathcal{L}\vec{v}_i = \lambda_i \vec{v}_i, \quad (4.96)$$

determine the critical properties of the Hamiltonian.

The eigenvalues λ_i are functions of b . Two successive transformations with scale change b make a total scale change b^2 , they are thus equivalent to a single transformation with the scale change b^2 .

$$\lambda_i(b)\lambda_i(b) = \lambda_i(b^2) \quad (4.97)$$

and the eigenvalues must be of the form

$$\lambda_i(b) = b^{x_i}. \quad (4.98)$$

x_i are the critical exponents and, as we shall see, they are closely related to the usual critical exponents α, β, \dots . So, if we know λ_i , we can obtain the critical exponents x_i from

$$x_i = \frac{\ln \lambda_i}{\ln b}. \quad (4.99)$$

- *Example 1:* For the $d = 1$ Ising model, $\mu_1 = x$ and $\mu_2 = y$. After linearization of the recursion relations (4.73) around the fixed point at $\vec{\mu}^* = (0, 1)$ we get:

$$\delta x' = 4\delta x \quad \delta y' = 2\delta y. \quad (4.100)$$

Hence, \mathcal{L} is diagonal in this representation. The eigenvalues and corresponding eigenvectors are:

$$\begin{aligned} \lambda_1 &= 4 & \nu_1 &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \lambda_2 &= 2 & \nu_2 &= \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \end{aligned} \quad (4.101)$$

It follows that the two critical exponents are both positive:

$$x_1 = 2 \quad x_2 = 1. \quad (4.102)$$

- *Example 2:* For the $d = 2$ Ising model, $\mu_1 = K$ and $\mu_2 = L$, and after linearization Eqs. (4.82) become

$$\begin{aligned} \delta K' &= 4K^* \delta K + \delta L \\ \delta L' &= 2K^* \delta K. \end{aligned} \quad (4.103)$$

For the non-trivial fixed point at $(K^*, L^*) = (\frac{1}{3}, \frac{1}{9})$, the RG transformation matrix is:

$$\mathcal{L} = \begin{vmatrix} 4K^* & 1 \\ 2K^* & 0 \end{vmatrix} = \begin{vmatrix} 4/3 & 1 \\ 2/3 & 0 \end{vmatrix} \quad (4.104)$$

The eigenvalues of \mathcal{L} are:

$$\lambda_1 = \frac{2 + \sqrt{10}}{3} \approx 1.72, \quad \lambda_2 = \frac{2 - \sqrt{10}}{3} \approx -0.387 \quad (4.105)$$

and the eigenvectors:

$$\begin{aligned} \vec{\nu}_1 &\propto \begin{pmatrix} 1 + \sqrt{10}/2 \\ 1 \end{pmatrix} \approx \begin{pmatrix} 2.58 \\ 1 \end{pmatrix} \\ \vec{\nu}_2 &\propto \begin{pmatrix} 1 - \sqrt{10}/2 \\ 1 \end{pmatrix} \approx \begin{pmatrix} -0.58 \\ 1 \end{pmatrix} \end{aligned} \quad (4.106)$$

(Question: what does negative λ_2 mean physically, how does the system behave under successive RG transformations if λ is negative?) The corresponding critical exponents are:

$$\begin{aligned} x_1 &= 2 \frac{\ln[(2 + \sqrt{10})/3]}{\ln 2} \approx 1.57 \\ x_2 &= 2 \frac{\ln|(2 - \sqrt{10})/3|}{\ln 2} \approx -2.74, \end{aligned} \quad (4.107)$$

one is positive and the other one is negative.

Flow in the parameter space. We first express $\delta\vec{\mu}$ near the fixed point at $\vec{\mu}^*$ in terms of the eigenvectors $\vec{\nu}_i$ of \mathcal{L} . In case of the $d = 1$ Ising model this means that

$$\begin{aligned} \delta x' &= \lambda_1 \delta x = 4\delta x \\ \delta y' &= \lambda_2 \delta y = 2\delta y. \end{aligned} \quad (4.108)$$

In general we write $\delta\vec{\mu}$ as:

$$\delta\vec{\mu} = \sum_i h_i \vec{\nu}_i. \quad (4.109)$$

The coefficients h_i are the *linear scaling fields*. They are the components of the generalized field in the coordinate system of the eigenvectors. Under a renormalization

$$\delta\vec{\mu}' = \mathcal{L}\delta\vec{\mu} = \sum_i h_i \mathcal{L}\vec{\nu}_i = \sum_i h_i \lambda_i \vec{\nu}_i, \quad (4.110)$$

where $\lambda_i = b^{x_i}$. On the other hand:

$$\delta\vec{\mu}' = \sum_i h'_i \vec{v}_i, \quad (4.111)$$

therefore

$$h'_i = b^{x_i} h_i \quad (b > 1). \quad (4.112)$$

It follows that the flow of the Hamiltonian in the parameter space close to the fixed point depends crucially on the exponents x_i . For positive x_i ($\lambda_i > 1$), the scaling field h_i increases under repeated RG transformations and drives the system away from the fixed point. In this case, h_i is a *relevant scaling field*. If the exponent x_i is negative, the corresponding scaling field h_i decreases under subsequent RG transformations, thus driving the system towards the fixed point. Such fields are called *irrelevant scaling fields*.

Thus the stability of the fixed point and the topology of the flow lines around it depend on the number of relevant and irrelevant scaling fields associated with it. To illustrate this, Fig. 4.18 shows the trajectories in the parameter space with two irrelevant and one relevant scaling field.

In general, any system which has initially the component of the relevant field different from zero, will after repeated RG transformations eventually be driven away from the fixed point (dotted lines in Fig. 4.18). Physically this corresponds to the system moving away from criticality.

The points in which the relevant field vanishes, define the *critical surface*. All the points on the critical surface flow *into* the fixed point. The points on the critical surface represent systems with an infinite correlation length. Under RG transformations, the systems stays on the critical surface (ξ remains infinite), but it moves towards the fixed point.

- *Example:* The $d = 1$ Ising model has $h_1 = x$ and $h_2 = y$. For the non-trivial fixed point at $x^* = 0$ and $y^* = 1$ the critical exponents are $x_1 = 2 > 0$ and $x_2 = 1 > 0$, that means that both, temperature and the external magnetic field, are relevant.

In the case of the $d = 2$ Ising model on the square lattice, $x_1 > 0$ and $x_2 < 0$, as we have seen. It has one relevant scaling field (h_1) and one irrelevant field (h_2). The relevant scaling field is related to reduced temperature whereas the irrelevant field is related to L , the coupling with NNN. Notice that here we considered the Ising model without external field. If the field

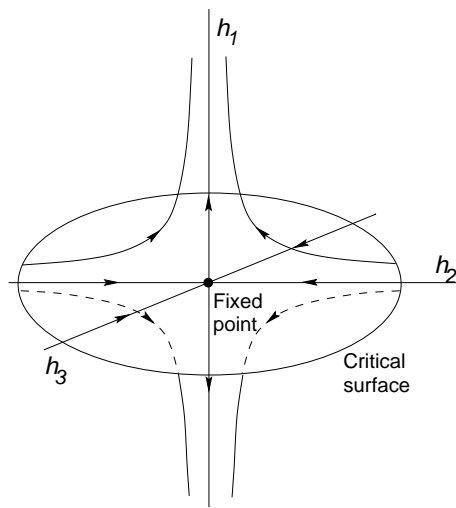


Figure 4.18: A fixed point with one relevant (h_1) and two irrelevant (h_2 and h_3) scaling fields. All the points on the two-dimensional critical surface flow to the fixed point (solid trajectories). The points slightly above (or below) the critical surface first flow towards the fixed point but eventually they are driven away from it as the relevant scaling field increases (dotted trajectories).

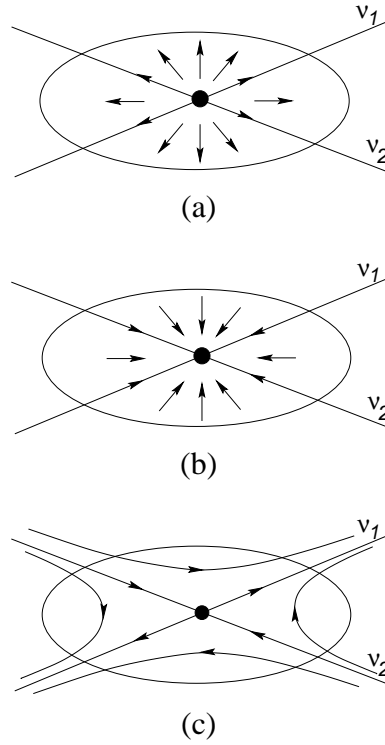


Figure 4.19: Topology of the RG flow near fixed points with: (a) two relevant fields, (b) two irrelevant fields, and (c) one relevant and one irrelevant field.

is included, then we have two relevant scaling fields and one irrelevant (the magnetic field is always relevant).

Figure 4.19 shows possible scenarios in the case of two scaling fields. Ferromagnets have two relevant fields, external field and temperature (or NN exchange interaction) and the fixed point is of the type shown in Fig. 4.19 (a).

Universality. The critical behaviour is completely determined by the behaviour of the system close to the fixed point, where the transformation matrix can be linearized. The transformation matrix depends on the values of the scaling fields at the critical point and not on the initial values of the irrelevant scaling fields. Thus, *all systems which flow close to the fixed point, independent of where in the parameter space they start, will exhibit the same critical exponents determined by the eigenvalues of the linear transformation matrix at that fixed point.* The critical

exponents are universal and do not depend on the initial point in the parameter space (like the coupling strength).

Critical exponent ν . To illustrate how the critical exponent ν is obtained from x_i , let us consider ferromagnetic systems. We know that there the critical point is specified by H and T . These are the relevant fields for magnetic systems. Therefore we identify h_1 by the reduced temperature, $h_1 = t$ and h_2 by the reduced magnetic field, $h_2 = h$. Under RG transformations ξ and t scale as

$$\xi' = \frac{\xi}{b} \quad t' = b^{x_t} t \quad (4.113)$$

The two equations tell us that $\xi^{x_t} t$ is constant, it is invariant under RG transformations. Therefore

$$\xi(t) \propto t^{-1/x_t} \quad (4.114)$$

Since, by definition of ν , $\xi \propto t^{-\nu}$, we get

$$\boxed{\nu = \frac{1}{x_t}} \quad (4.115)$$

The last equation can be obtained also in a different way. Unless the system is initially at the critical point, the correlation length will decrease under successive block-spin RG transformations. The system will eventually reach a point where the correlation length ξ' will become equal to the size of the renormalized unit cell, so that the nearest-neighbour spins become uncorrelated. This point corresponds to a certain temperature t_0 which cannot depend on the initial temperature. Suppose, the system is initially at the temperature t , and arrives at t_0 after n iterations. Then,

$$t_0 = b^{n x_t} t \quad \text{and} \quad \xi' \equiv 1 = \xi / b^n \Rightarrow \xi = b^n \quad (4.116)$$

From these equations we get Eq. 4.115 again.

- We return again to the $d = 2$ Ising model. Above, we considered the model at $h = 0$. Thus, the exponent x_t corresponds to the relevant exponent $x_1 = 1.57$ from Eq. (4.107) and therefore $\nu = 1/x_1 = 0.64$. This is better than the mean-field value of $1/2$, but still far from the exact value $\nu = 1$. The discrepancy has its origin in the approximations we have made to obtain the RG recursion relations.

Other critical exponents. The original and renormalized Hamiltonians \mathcal{H} and \mathcal{H}' correspond to positions in the parameter space, $\vec{\mu}$ and $\vec{\mu}'$. We apply relation (4.86) to the *singular part* of the reduced free energy per spin:

$$g_s(\vec{\mu}') = b^d g_s(\vec{\mu}). \quad (4.117)$$

Near a fixed point $\vec{\mu}$ and $\vec{\mu}'$ can be expressed in terms of the scaling fields h_i and h'_i which are related by (4.112). Hence eq. (4.117) becomes:

$$g_s(h_1, h_2, \dots) = b^{-d} g_s(b^{x_1} h_1, b^{x_2} h_2, \dots), \quad (4.118)$$

where g_s is a homogeneous function of $\vec{\mu}$. This is the *scaling form* of the singular part of the free energy. This equation holds for arbitrary b because we can repeat transformations many times. In fact, b can be any real number > 1 ! Again we identify $h_1 = t$, $h_2 = h$ and assume that other scaling fields are irrelevant. Then, eq. (4.118) becomes

$$g_s(t, h) = b^{-d} g_s(b^{x_1} t, b^{x_2} h). \quad (4.119)$$

Now, to come to the specific-heat critical exponent α , we remind that

$$c_H = \left(\frac{\partial^2 g_s}{\partial t^2} \right)_{h=0} \equiv g_{tt}(h=0) \propto |t|^{-\alpha}. \quad (4.120)$$

Differentiating Eq. (4.119) twice with respect to temperature and putting the scaling field h equal to zero gives:

$$g_{tt}(t, h=0) \propto b^{-d} g_{\tau\tau}(\tau, 0) \frac{d\tau}{dt} = b^{-d+2x_1} g_{\tau\tau}, \quad (4.121)$$

where we have used $\tau = b^{x_1} t$. Initially, t is small whereas $b > 1$ so that the product $b^{x_1} t$ increases with each iteration. We repeat the iterations until

$$b^{x_1} |t| = 1. \quad (4.122)$$

In this way, the arguments of $g_{\tau\tau}$ and therefore also the function $g_{\tau\tau}$ are constant, all the temperature dependence of the specific heat is in the prefactor b^{-d+2x_1} :

$$g_{tt}(t, h=0) \propto |t|^{(d-2x_1)/x_1} g_{tt}(\pm 1, 0) \quad (4.123)$$

and

$$\alpha = 2 - \frac{d}{x_1}. \quad (4.124)$$

Similarly we find

$$\begin{aligned}\beta &= \frac{d - x_2}{x_1} \\ \gamma &= \frac{2x_2 - d}{x_1} \\ \delta &= \frac{x_2}{d - x_2}\end{aligned}\tag{4.125}$$

We have expressed five critical exponents in terms of two independent exponents x_1 and x_2 . That means that there must be relations between the exponents $\alpha \cdots \nu$. One can easily verify the scaling laws:

$$\begin{aligned}\alpha + 2\beta + \gamma &= 2 \\ \gamma &= \beta(\delta - 1) \\ 2 - \alpha &= d\nu.\end{aligned}\tag{4.126}$$

4.4 Momentum-Space Renormalization Group

RG methods in momentum space are much more powerful and more widely used than their counterparts in real space. The problem is only that they are much more complicated than the above real-space prototype examples. We shall define the RG transformation for the Landau model, that means that we shall now use continuous instead of discrete spin variables. The main reason is that *discrete spins become essentially continuous after (block-spin) transformations*.

We start with the free energy of the form:

$$\mathcal{H}(h, m(r)) = \int_V d^d \vec{r} \left[\frac{1}{2} g |\nabla m(r)|^2 - m(r) h + \frac{1}{2} r m^2(r) + u m^4(r) \right]\tag{4.127}$$

This is also called the Landau-Ginzburg-Wilson (LGW) Hamiltonian. The main procedure of all the RG transformations is "coarse graining" (i.e., transformation to larger unit cells) followed by rescaling (to make the system look like the original one). In real-space RG this was made either by introducing block spins or by decimation. Now we shall make coarse graining in *momentum space*. Instead

of *expanding* the size of the unit cell in real space, we shall *shrink* the size of the Brillouin zone in the momentum space. In the following the method of the momentum-space renormalization group will be demonstrated on the Gaussian model.

4.4.1 The Gaussian Model

The Hamiltonian of the Gaussian model in the momentum space is:

$$\mathcal{H} = \frac{1}{2} \sum_{|\vec{q}| < \Lambda} (r + gq^2) |m(\vec{q})|^2 - hm(0). \quad (4.128)$$

We assume a homogeneous magnetic field and we are only interested in the stability of the $m = 0$ solution ($r > 0$), therefore we take $u = 0$; this will simplify the summation over \vec{q} . Λ is the cutoff momentum (Brillouin zone boundary). Taking $u = 0$ means that the modes with different momenta become uncoupled.

The corresponding partition function is equal to the functional integral over all functions $m(r)$. In momentum space, the functional integral becomes:

$$\int \mathcal{D}m(r) \rightarrow \prod_{|\vec{q}| < \Lambda} \int_0^\infty dm(q) \quad (4.129)$$

and the partition function:

$$Y(t, h) = \prod_{|\vec{q}| < \Lambda} \int_0^\infty dm(q) e^{-\frac{1}{2} \sum_{|\vec{q}| < \Lambda} (r + gq^2) |m(q)|^2 - hm(0)}. \quad (4.130)$$

The RG transformation consists of "integrating out" the wavevectors that are outside the sphere of the radius Λ/b and then rescaling. One RG cycle consists of three steps:

1. **Integration.** After "integrating out" the momenta within a thin shell between the radii Λ/b and Λ , the partition function (which should not change) becomes:

$$Y = e^\Omega \prod_{|\vec{q}| < \Lambda/b} \int dm(q) e^{-\mathcal{H}'}, \quad (4.131)$$

where the first term, $\exp(\Omega)$, is a function of Λ , b , and coupling constants. It comes from the integration over momenta in the shell between Λ/b and Λ .

The second term is the partition function of the states with momenta $|q| < \Lambda/b$. Notice that this separation into two contributions was only possible because we set $u = 0$, the partition function is then a product of independent Gaussian integrals. The new Hamiltonian,

$$\mathcal{H}'(m(q)) = \sum_{|q| < \Lambda/b} \frac{1}{2}(r + gq^2)|m(q)|^2 - hm(0), \quad (4.132)$$

depends only on $m(q)$ with $|q| < \Lambda/b$ and has the same structure as the original one.

2. **Rescaling.** The second step restores the old cutoff by blowing up the radius of integration to the original value Λ in such a way that we change the variable of integration to

$$\vec{q}' = b\vec{q}. \quad (4.133)$$

In this way the cutoff momentum is restored back to Λ . This step corresponds to putting $a' = ba$ in real-space RG. The Hamiltonian now reads:

$$\mathcal{H}'(m) = b^{-d} \sum_{|q'| < \Lambda} \frac{1}{2}(r + g\frac{q'^2}{b^2})|m(\frac{q'}{b})|^2 - hm(0). \quad (4.134)$$

3. **Normalization.** The step 3 restores the standard normalization of the order parameter. This step corresponds to putting the new spin equal to 1 after block-spin or decimation in real-space RG. Renormalization is restored with $g' = g$ if we set:

$$m'(q') = \sqrt{\frac{1}{b^{d+2}}} m(\frac{q'}{b}). \quad (4.135)$$

The transformed Hamiltonian and the partition function (which must not change) now look like

$$\mathcal{H}'(m) = \sum_{|q'| < \Lambda} \frac{1}{2}(r' + gq'^2)|m'(q')|^2 - h'm'(0). \quad (4.136)$$

and

$$Y = e^{\Omega} \prod_{|q'| < \Lambda} \int dm(q') e^{-\beta\mathcal{H}'}. \quad (4.137)$$

In this way we integrated out large momenta and transformed \mathcal{H} to its original form. The recursion relations are:

$$\begin{aligned} r' &= b^2 r \\ h' &= b^{(d+2)/2} h. \end{aligned} \quad (4.138)$$

Since $b \approx 1$ we can write $b = 1 + d\tau$, thus,

$$b^n \approx 1 + n \ln b \equiv 1 + n d\tau. \quad (4.139)$$

We can say that τ counts the number of RG cycles. Then, r and h obey the following differential equations:

$$\begin{aligned} \frac{dr}{d\tau} &= 2r \\ \frac{dh}{d\tau} &= \frac{d+2}{2} h \end{aligned} \quad (4.140)$$

with the solution

$$\begin{aligned} r &= r_0 e^{x_t \tau} & x_t &= 2 \\ h &= h_0 e^{x_h \tau} & x_h &= \frac{d+2}{2}. \end{aligned} \quad (4.141)$$

Since q is continuous (for macroscopic systems), the RG transformations was carried out in infinitesimal steps ($b \approx 1$) and the problem was reduced to solving a set of differential equations.

The Gaussian model has two trivial fixed points at $h^* = 0$, one at $r^* = \infty$ (very high T , $m = 0$) and one at $r^* = -\infty$ (very low T , where $m \rightarrow \infty$ because we set $u = 0!$), and one nontrivial fixed point (the critical point) at $h^* = r^* = 0$. At the nontrivial fixed point the critical exponents are both positive and r and h are both relevant scaling fields. From x_i , we shall calculate other critical exponents in the Section 4.4.3.

As we shall see in the next Section, for $d < 4$ the Gaussian model is unstable with respect to the parameter u coming from the four-spin interaction in the Hamiltonian.

4.4.2 The Landau-Wilson Model

In the Gaussian model, we have neglected the fourth and higher-order terms in the Landau Hamiltonian. The next step is to include also the m^4 term. This will lead to non-MFA exponents, as we shall see. Here we shall bring only an outline of the RG procedure and discuss the results. We start with the LGW Hamiltonian in real space,

$$\mathcal{H}(m) = \int d^d \vec{r} \left[\frac{1}{2} g |\nabla m(\vec{r})|^2 + \frac{1}{2} r m^2(\vec{r}) + u m^4(\vec{r}) \right] \quad (4.142)$$

where $u > 0$. We first integrate out an infinitesimally thin shell of the thickness δq in momentum space by taking $b \approx 1$. Then,

$$\delta q / \Lambda = \delta \Lambda / \Lambda = \delta \ln \Lambda = b - 1 \approx \ln b \quad (4.143)$$

and we write

$$m(\vec{r}) = \bar{m}(\vec{r}) + \delta m(\vec{r}) \quad (4.144)$$

where $\delta m(\vec{r})$ contains the Fourier components to be integrated out and \bar{m} the rest. After some approximations we get the following renormalization-group (differential) equations for r and u [Huang, 1987]:

$$\begin{aligned} \frac{dr(\tau)}{d\tau} &= 2r + 12\Lambda^{d-2}u \left(1 - \frac{r}{\Lambda^2}\right) \\ \frac{du(\tau)}{d\tau} &= (4-d)u - 36\Lambda^{d-4}u^2. \end{aligned} \quad (4.145)$$

To analyse these equations, it is convenient to introduce the dimensionless coupling constants:

$$x = \frac{r}{\Lambda^2}, \quad y = \frac{u}{\Lambda^{4-d}} \quad (4.146)$$

then, the RG equations become:

$$\begin{aligned} \frac{dx}{d\tau} &= 2x + 12y(1-x) \\ \frac{dy}{d\tau} &= \epsilon y - 36y^2 \end{aligned} \quad (4.147)$$

where

$$\epsilon \equiv 4 - d \quad (4.148)$$

We shall assume that ϵ is small. The system has two fixed points at finite x and y , fixed by $dx/d\tau = dy/d\tau = 0$:

$$\left. \begin{array}{l} x^* = 0 \\ y^* = 0 \end{array} \right\} \quad \text{Gaussian fixed point}$$

$$\left. \begin{array}{l} x^* = -\epsilon/6 \\ y^* = \epsilon/36 \end{array} \right\} \quad \text{“Nontrivial” fixed point.} \quad (4.149)$$

The Gaussian fixed point is at $x = y = 0$ whereas the “nontrivial” (also called the “Wilson–Fisher”) fixed point lies in the upper ($y > 0$) half of the (x, y) plane only if $\epsilon > 0$, $d < 4$. In the lower half of the (x, y) plane, u is negative and the system is unstable. The nontrivial fixed point approaches the Gaussian fixed point as $\epsilon \rightarrow 0$.

In the neighbourhood of the fixed points we linearize the RG equations

$$x = x^* + \delta x, \quad y = y^* + \delta y \quad (4.150)$$

and get the following linearized recursion relations:

$$\begin{aligned} \frac{d}{d\tau} \delta x &= (2 - 12y^*)\delta x + 12(1 - x^*)\delta y \\ \frac{d}{d\tau} \delta y &= (\epsilon - 72y^*)\delta y \end{aligned} \quad (4.151)$$

which are written in the general form as:

$$\frac{d}{d\tau} \begin{pmatrix} \delta x \\ \delta y \end{pmatrix} = \mathcal{L} \begin{pmatrix} \delta x \\ \delta y \end{pmatrix} \quad (4.152)$$

At the Gaussian fixed point the linear matrix \mathcal{L} is:

$$\mathcal{L} = \begin{vmatrix} 2 & 12 \\ 0 & \epsilon \end{vmatrix} \quad (4.153)$$

The matrix \mathcal{L} has the following eigenvalues and eigenvectors:

$$\lambda_1 = 2 \quad \vec{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\lambda_2 = \epsilon \quad \vec{\nu}_2 \propto \begin{pmatrix} 1 \\ -(2 - \epsilon)/12 \end{pmatrix} \quad (4.154)$$

The corresponding scaling fields are h_1 and h_2 . Under the (infinitesimal) RG transformations the scaling fields change as $dh_i/d\tau = \lambda_i h_i$,

$$h_i(\tau) = h_{i0} e^{\lambda_i \tau} = h_{i0} b^{\lambda_i} \quad (4.155)$$

Thus, the critical exponents are equal to the eigenvalues λ_i : $x_t = \lambda_1 = 2$ and $x_2 = \lambda_2 = \epsilon$

In the neighbourhood of the nontrivial fixed point the linear transformation matrix \mathcal{L} is:

$$\mathcal{L} = \begin{vmatrix} 2(1 - 6y^*) & 12(1 - x^*) \\ 0 & \epsilon - 72y^* \end{vmatrix} = \begin{vmatrix} 2(1 - \epsilon/6) & 12(1 + \epsilon/6) \\ 0 & -\epsilon \end{vmatrix} \quad (4.156)$$

It has the following eigenvalues and eigenvectors:

$$\begin{aligned} \lambda_1 &= 2 - \epsilon/3 & \nu_1 &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \lambda_2 &= -\epsilon & \nu_2 &= \begin{pmatrix} -\frac{6+\epsilon}{1+\epsilon/3} \\ 1 \end{pmatrix} \end{aligned} \quad (4.157)$$

Now we have the following picture. For $d < 4$ ($\epsilon > 0$) the nontrivial fixed point lies in the upper half of the (x, y) plane whereas the Gaussian fixed point is (as before) at $x = y = 0$, see Fig. 4.20. For the nontrivial fixed point, h_2 is irrelevant (it does not affect the critical exponents) and $h_1 = r \propto t$ the relevant scaling field whereas for the Gaussian fixed point, both fields are relevant. Thus, the line ν_2 is the critical line. For $d > 4$ ($\epsilon < 0$) the nontrivial fixed point lies in the lower half of the (x, y) plane, i.e., in the unphysical region ($u < 0$), see Fig. 4.21, and the critical behaviour is governed by the Gaussian fixed point. $d = 4$ is the special case when both fixed points coincide. The linearized recursion relations give $x_2 = 0$, h_2 is marginal, and x_2 cannot tell us the direction of RG flow. In this case we must include higher-order terms which were neglected when we linearized the transformation matrix. These terms tell us that the scaling field h_2 is irrelevant for $u > 0$ at $d = 4$.

4.4.3 Critical Exponents

We have seen that the critical exponents of the nontrivial fixed point are

$$x_t = 2 - \epsilon/3$$

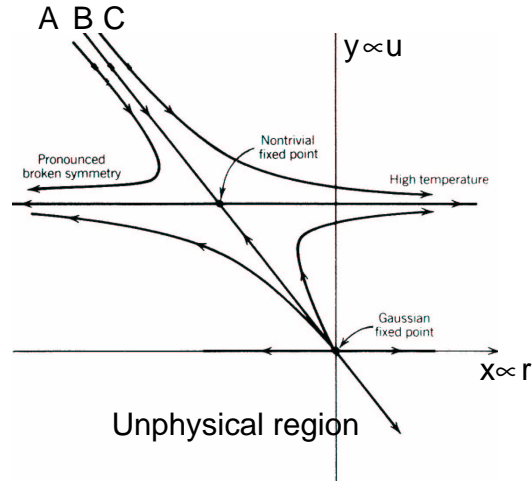


Figure 4.20: Fixed points and flow diagram of the Landau-Ginzburg-Wilson model for $d < 4$. The Gaussian fixed point at $(x^*, y^*) = (0, 0)$ has the mean-field exponents. The non-trivial fixed point at $(x^*, y^*) = (-\epsilon/6, \epsilon/36)$ is unstable against the relevant perturbation $\delta x \sim t$ and stable against the irrelevant perturbation δy . The system at A has $t < 0$ and it flows towards the low-temperature fixed point, r becomes increasingly negative. At the point C the system has $t > 0$ and it flows to the high-temperature fixed point, r becomes increasingly positive. The point B is on the critical line and the system flows towards the non-trivial fixed point. (After K. Huang, Statistical Mechanics)

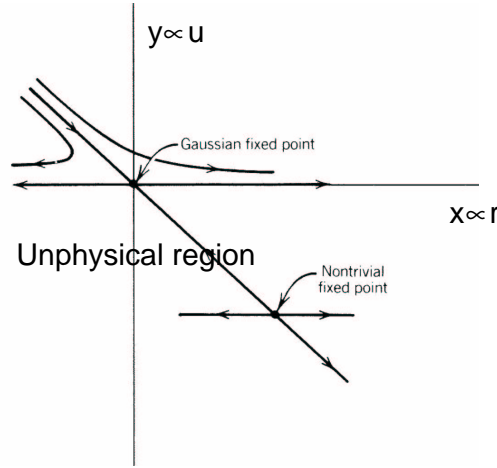


Figure 4.21: Fixed points and flow diagram of the Landau-Ginzburg-Wilson model for $d > 4$. The nontrivial fixed point is in the unphysical region. To stabilize the system in the lower half of the (x, y) plane, one needs an m^6 term. (After K. Huang, Statistical Mechanics)

$$\begin{aligned} x_u &= -\epsilon \\ x_h &= 1 + d/2. \end{aligned} \quad (4.158)$$

In fact, we calculated x_h only for the Gaussian fixed point in the Eq. (4.138). However, the term with m^4 does not affect the value of x_h so that x_h of the nontrivial fixed point is the same. With these values of x_i we get from Eqs. (4.115, 4.124, 4.125) to the first order in ϵ for $d < 4$:

$$\begin{aligned} \alpha &= \frac{\epsilon}{6} \\ \beta &= \frac{1}{2} - \frac{\epsilon}{6} \\ \gamma &= 1 + \frac{\epsilon}{6} \\ \delta &= 3 + \epsilon \\ \nu &= \frac{1}{2} + \frac{\epsilon}{12} \\ \eta &= 0. \end{aligned} \quad (4.159)$$

For $d \geq 4$ ($\epsilon \leq 0$) we only have the Gaussian fixed point in the physical region ($u > 0$) with the exponents

$$\begin{aligned} x_t &= 2 \\ x_u &= \epsilon \\ x_h &= 1 + d/2. \end{aligned} \tag{4.160}$$

With Eqs. (4.115, 4.124, 4.125) we find:

$$\begin{aligned} \alpha &= 2 - \frac{d}{2} \\ (\beta &= \frac{1}{2} - \frac{\epsilon}{4}) \\ \gamma &= 1 \\ (\delta &= 3 + \epsilon) \\ \nu &= \frac{1}{2} \\ \eta &= 0. \end{aligned} \tag{4.161}$$

The above expressions for β and δ are wrong, because we have neglected the m^4 term in considering the Gaussian fixed point. Below T_C , r is negative and for m to be finite, the term um^4 with $u > 0$ must be included in the Hamiltonian. Although u is an irrelevant field, it helps to stabilize m at a finite value and in this way influences the critical behaviour. By treating the um^4 term in the Gaussian approximation (independent modes) we get the usual mean-field values for the exponents β and η :

$$\begin{aligned} \beta &= \frac{1}{2} \\ \delta &= 3 \end{aligned} \tag{4.162}$$

for all $d \geq 4$. These results are correct and also agree with the results of the Landau (MFA) model in Sec. 4.1.4.

The Gaussian model gives the same critical exponents as the mean-field theory except for the specific-heat exponent α . Therefore the Gaussian fixed point is identified as the mean-field fixed point. In the Gaussian model the specific heat

diverges ($\alpha > 0$) for $d < 4$. This divergence is caused by the long-wavelength fluctuations which were neglected in the calculation of the free energy in the MF approximation.

In $d = 3$ ($\epsilon = 1$) we find $\alpha \approx 0.17$ and $\nu \approx 0.58$. The results of this paragraph are based on an expansion of d around $d = 4$ and are valid for small $\epsilon > 0$ or for $d \geq 4$. Nevertheless, we put $\epsilon = 1$ when $d = 3$. That means that ϵ is not small and higher-order terms in ϵ must be included. This is then the ϵ -expansion, which gives good values for the critical exponents.

Further Reading

- [1] K. Binder (ed.), *Monte-Carlo methods in statistical physics*, Springer-Verlag (New York, 1986).
- [2] J.J. Binney, N.J. Dowrick, A.J. Fisher, and M.E. Newman, *The Theory of Critical Phenomena*, Oxford University Press (Oxford, 1992).
- [3] J. Cardy, *Scaling and Renormalization in Statistical Physics*, Cambridge University Press (Cambridge, 1996).
- [4] W. Greiner, L. Neise, and H. Stöcker, *Thermodynamics and Statistical Mechanics*, Springer-Verlag (New York, 1995).
- [5] K. Huang, *Statistical Mechanics*, J. Wiley (New York, 1987).
- [6] R. Kubo, *Statistical Mechanics*, North Holland (Amsterdam, 1967).
- [7] L.D. Landau and E.M. Lifshitz, *Statistical Physics I,II*, Pergamon Press (Oxford, 1980).
- [8] S.-k. Ma, *Modern Theory of Critical Phenomena*, Benjamin (Reading, 1976).
- [9] B.M. McCoy and T.T. Wu, *The Two-Dimensional Ising Model*, Harvard (Cambridge, 1973).
- [10] M. Plischke and B. Bergersen, *Equilibrium Statistical Mechanics*, World Scientific (Singapore, 1994).
- [11] F. Schwabl, *Statistical Mechanics*, Springer Verlag (Berlin, 2002).
- [12] H.E. Stanley, *Phase Transitions and Critical Phenomena*, Oxford University Press (London, 1971).

- [13] K.G. Wilson, *Scientific American*, **241**, 140 (1979).
- [14] Julia M Yeomans, *Statistical Mechanics of Phase Transitions*, Clarendon (Oxford, 1992).

Appendix A

Distribution Functions

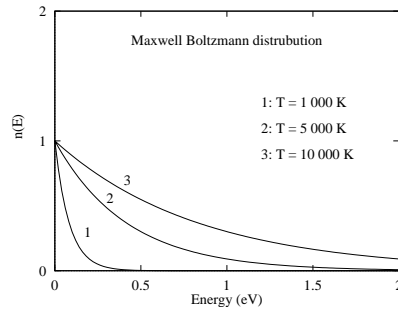
We have to distinguish three kinds of systems:

- (a) identical but distinguishable particles ("classical particles")
- (b) identical indistinguishable particles with half-integer spin (Fermi-ions)
- (c) identical indistinguishable particles with integer spins (Bosons)

The occupation probability of a state i with energy E_i is:

(a-Maxwell-Boltzmann):

$$n(E_i) = \frac{1}{e^{\beta E_i}} \quad (\text{A.1})$$

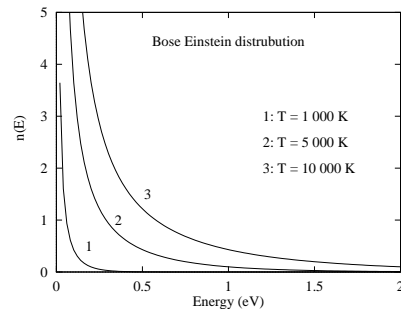
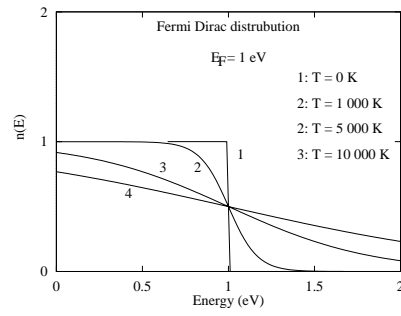


(b-Fermi-Dirac):

$$n(E_i) = \frac{1}{e^{\beta(E_i - \varepsilon_F)} + 1} \quad (\text{A.2})$$

(c-Bose-Einstein):

$$n(E_i) = \frac{1}{e^{\beta E_i} - 1}. \quad (\text{A.3})$$



Appendix B

Maxwell Relations

Let df be a total differential of the function $f(x, y)$. Then,

$$df = u(x, y)dx + v(x, y)dy \quad (\text{B.1})$$

with

$$u(x, y) = \left(\frac{\partial f}{\partial x} \right)_y; \quad v(x, y) = \left(\frac{\partial f}{\partial y} \right)_x \quad (\text{B.2})$$

and

$$\left(\frac{\partial u}{\partial y} \right)_x = \left(\frac{\partial^2 f}{\partial y \partial x} \right) = \left(\frac{\partial^2 f}{\partial x \partial y} \right) = \left(\frac{\partial v}{\partial x} \right)_y. \quad (\text{B.3})$$

This is a very useful relation that - when applied to thermodynamic functions - gives the *Maxwell relations*:

$$\left(\frac{\partial T}{\partial V} \right)_S = - \left(\frac{\partial p}{\partial S} \right)_V, \quad (\text{B.4})$$

$$\left(\frac{\partial T}{\partial p} \right)_S = \left(\frac{\partial V}{\partial S} \right)_p, \quad (\text{B.5})$$

$$\left(\frac{\partial S}{\partial p} \right)_T = - \left(\frac{\partial V}{\partial T} \right)_p, \quad (\text{B.6})$$

$$\left(\frac{\partial S}{\partial V} \right)_T = \left(\frac{\partial p}{\partial T} \right)_V. \quad (\text{B.7})$$

Appendix C

Basic Thermodynamic Relations of Magnetic Systems Revisited

The *internal energy* is:

$$U(\mathcal{S}, M) = - \sum_{\langle ij \rangle} J_{i,j} \langle S_i S_j \rangle \quad (\text{C.1})$$

where \mathcal{S} is the entropy, M is the total magnetic moment (in dimensionless units), $M = Nm$, m is the magnetic moment per lattice site, $m = \langle S_i \rangle$. $J_{i,j}$ is the exchange interaction between the spins i and j . The (magnetic) *enthalpy* is (here we will use E for the enthalpy, to avoid confusion with the field H):

$$E(\mathcal{S}, H) = U - HM \equiv \langle \mathcal{H} \rangle \quad (\text{C.2})$$

where H is the magnetic field in units of energy. The *Helmholtz* free energy is:

$$F(T, M) = -k_B T \ln Z(T, M, N) = U - T\mathcal{S} \quad (\text{C.3})$$

and its total differential is

$$dF = -\mathcal{S} dT + H dM \quad (\text{C.4})$$

(see Table II of Part I; the number of spins N is constant): The *Gibbs* free energy is

$$\begin{aligned} G(T, H) &= -k_B T \ln Y(T, H) = E - T\mathcal{S} \\ &= F - HM = U - T\mathcal{S} - HM \end{aligned} \quad (\text{C.5})$$

and its differential is

$$dG = -S dT - M dH \quad (\text{C.6})$$

Here, we introduced also the corresponding partition functions Z and Y . From Eq. (C.4) we get immediately the *equation of state*:

$$H = \left(\frac{\partial F}{\partial M} \right)_{T,N} \quad (\text{C.7})$$

From Eq. (C.6) we also get an equation of state:

$$M = - \left(\frac{\partial G}{\partial H} \right)_{T,N} = - \left(\frac{\partial E}{\partial H} \right)_{S,N} \quad (\text{C.8})$$

Of course, the equations (C.7) and (C.8) must lead to the same equation, they are just two ways of writing the equation of state. When we want to study systems at constant H (and not at constant M), the more convenient thermodynamic potential is the *Gibbs free energy* $G(T, H, N)$ instead of the Helmholtz free energy.

Other useful thermodynamic relations are:

$$C_H = T \left(\frac{\partial S}{\partial T} \right)_H = \frac{dU}{dT} = -T \left(\frac{\partial^2 G}{\partial T^2} \right)_H \quad (\text{C.9})$$

$$C_M = T \left(\frac{\partial S}{\partial T} \right)_M = -T \left(\frac{\partial^2 F}{\partial T^2} \right)_H \quad (\text{C.10})$$

$$\chi_T = \left(\frac{\partial M}{\partial H} \right)_T = - \left(\frac{\partial^2 G}{\partial H^2} \right)_T \quad (\text{C.11})$$

$$\chi_S = \left(\frac{\partial M}{\partial H} \right)_S = - \left(\frac{\partial^2 E}{\partial H^2} \right)_S \quad (\text{C.12})$$